BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

PATENT
Docket No.: **19603/481 (CRF D-2472A)**

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s)	:	Barany et al.	Examiner: B. J. Forman
Serial No.	:	09/528,014	
Cnfrm. No.	:	4478	Art Unit: 1634
Filed	:	March 17, 2000	
For	:	COUPLED POLYMERASE CHAIN REACTION-RESTRICTION ENDONUCLEASE DIGESTION-LIGASE DETECTION REACTION PROCESS	

REQUEST FOR RECONSIDERATION

Mail Stop RCE
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

In response to the November 19, 2002, outstanding office action, reconsideration is respectfully requested.

The rejection of claims 1, 3-7, and 17 under 35 U.S.C. § 103(a) for obviousness over WO 97/31256 to Barany et al., ("Barany") in view of Jacobson et al., "A Highly Sensitive Assay for Mutant *ras* Genes and its Application to the Study of Presentation and Relapse Genotypes in Acute Leukemia," *Oncogene* 9(2):553-563 (1994) ("Jacobson") is respectfully traversed.

Barany teaches a method for identifying one or more of a plurality of sequences differing by one or more single base changes, insertions, deletions, or translocations in a plurality of target nucleotide sequences. The method includes a ligation phase, a capture phase, and a detection phase. The ligation phase utilizes a ligation detection reaction ("LDR") between one oligonucleotide probe having a target sequence-specific portion and an addressable array-specific portion, and a second oligonucleotide probe having a target sequence-specific portion and a detectable label. After the ligation phase, the capture phase is carried out by hybridizing the ligated oligonucleotide probes to a solid support with

an array of immobilized capture oligonucleotides, at least some of which are complementary to the addressable array-specific portion. Following completion of the capture phase, a detection phase is carried out to detect the labels of ligated oligonucleotide probes hybridized to the solid support. The ligation phase can be preceded by an amplification process (e.g., polymerase chain reaction ("PCR")). Barany does not teach the method wherein after the PCR, a second PCR is performed to produce secondary PCR products comprising a restriction enzyme site followed by restriction digest and third PCR step.

Jacobson teaches a method called PCR-Primer Introduced Restriction with Enrichment of Mutant Alleles ("PCR-PIREMA") for the detection of mutant *ras* genes. In PCR-PIREMA, the starting sample is amplified by the polymerase chain reaction ("PCR") technique using different "matched primers" to specifically amplify one of the known sites of mutations in a *ras* gene. A second PCR is carried out on the primary PCR extension product with mismatched "screening" primers to create a restriction endonuclease ("RE") site in the secondary extension product. The formation of the restriction site is dependent upon both introduced substitution and the normal codon sequence. Therefore, only normal alleles contain the introduced restriction site. The secondary extension products are digested with a restriction endonuclease recognizing the newly-created RE site, thereby reducing the presence of full-length normal sequences in the sample. Next, enrichment of the mutant allele population is carried out by subjecting the sample to a third PCR using the "screening primer" of the second PCR step, followed by a second restriction digestion, to "enrich" the sample for mutant alleles. The sample is subjected to gel electrophoresis, and the full-length (digestion-resistant) mutant bands are gel-purified. In an unspecified number of reactions, an "additional enrichment" step consisting of a third PCR/RE reaction and digest is carried out after the gel purification step. The digestion-resistant band indicates the presence of a mutant allele, but does not define the sequence of a mutant allele. To demonstrate that a mutant allele was present originally, the gel-purified band is further amplified and digested using the same "screening" primer and restriction enzyme as used in the secondary PCR step (i.e., "enriched screening). To define the mutation, yet another PCR is performed on the gel-purified product using mismatched "verification" primers that introduce new restriction sites into the PCR products derived from specific mutant alleles, the opposite of the screening strategy. The presence of a PCR product that cuts with the verification enzyme confirms that a specific mutation is present.

Nowhere does Jacobson teach using the ligase detection reaction in conjunction with its above-described procedure. Nowhere does Jacobson teach or suggest that the steps of the method described therein are useful or capable of being incorporated into another detection methodology.

A proper *prima facie* showing of obviousness requires the U.S. Patent and Trademark Office ("PTO") to satisfy three requirements. First, the prior art relied upon, coupled with knowledge generally available to one of ordinary skill in the art, must contain some suggestion which would have motivated the skilled artisan to combine or modify references. *In re Fine*, 837 F.2d 1071, 1074, 5 USPQ2d 1596, 1598 (Fed. Cir. 1988). Second, the PTO must show that, at the time the invention was made, the proposed modification had a reasonable expectation of success. *Amgen v. Chugai Pharm. Co.*, 927 F.2d 1200, 1208-1209, 18 USPQ2d 1016, 1023 (Fed. Cir. 1991). Finally, the combination of references must teach or suggest each and every limitation of the claimed invention. *In re Wilson*, 424 F.2d 1382, 1385, 165 USPQ 494, 496 (CCPA 1970).

The PTO fails, first of all, to show that the references cited, coupled with the knowledge generally available to one of ordinary skill in the art, contain any suggestion or motivation to combine or modify the references. Moreover, as demonstrated below, based on the accompanying Declaration of Joseph P. Day, Ph.D. Under 37 C.F.R. § 1.132 ("Day Declaration"), as well as the exhibits referenced therein, it would *not* have been obvious to one skilled in the art to combine the method of Jacobson with a method for identifying one or more low abundance sequences differing by one or more single-base changes in a sample containing a plurality of target nucleotide sequences.

A scientist, considering the use of Jacobson's technique in conjunction with a method for identifying one or more low abundance sequences differing by one or more single-base changes in a sample containing a plurality of target nucleotide sequences, would need to know whether the use of that technique achieved an increase in detection sensitivity, and if so, whether such increased sensitivity had a detrimental effect (Day Declaration ¶ 6). As described in greater detail below, the method of Jacobson has a serious shortcoming, in particular, a high potential for false positive results, which a skilled scientist would have recognized (*Id.*). Accordingly, there would be no reason to use Jacobson's technique in conjunction with a method for identifying one or more low abundance sequences differing by one or more single-base changes in a sample containing a plurality of target nucleotide sequences (*Id.*).

PCR-PIREMA is described by Jacobson as an "extremely sensitive" method of detecting mutant alleles in a biological sample having both normal and, potentially, mutant alleles of a particular DNA (Day Declaration ¶ 7). However, what a skilled scientist would have appreciated about Jacobson is that in order to reach a detection level of one mutant allele in 1,000,000 normal alleles (1 in 10^6) that Jacobson asserts PCR-PIREMA is capable of achieving, a single PCR-PIREMA sample, as described above, is subjected to PCR no fewer than three times, at least two RE digestion steps, one or more gel electrophoresis steps, and one or more gel purification procedures (*Id.*). These multiple steps make the method of Jacobson unattractive for several reasons (*Id.*).

Firstly, as described in the instant application (pg. 7, line 9 to pg. 8, line 23) and by others, the method of Jacobson is an adaptation of the commonly used PCR-based gene analysis technique known as PCR-RFLP (analysis by generation of artificial restriction fragment length polymorphisms), which requires repetitive steps to ensure an enrichment for allelic variants (Day Declaration ¶ 8). Although PCR-RFLP-based protocols achieve a relatively high sensitivity for mutant alleles, up to 1 in 10^4, the literature teaches that this method bears an increasing risk of polymerase-borne infidelity as the error rate of *Taq* or other thermostable polymerases is surpassed following repeated cycling (*Id.*). Errors due to infidelity have been found to be introduced early in the PCR process. Therefore, when the total number of PCR cycles involved exceeds approximately 70-80 cycles, as in Jacobson, the likelihood of error is significant (*Id.*). The rate of misincorporation is not a problem in most PCR systems (*Id.*). However, when the concentration of the desired DNA is low in the starting sample, as is the case in the PCR-PIREMA protocol of Jacobson, misincorporation is a serious problem, as acknowledged by Jacobson:

> In theory, another potential pitfall to this method is that after enrichment, a digestion resistant band could contain PCR products which had misincorporated a base during 1° or 2° PCR; such products might arise because *Taq* polymerase has a misincorporation rate estimated at 1/6000-1/8000 bases. This rate does not affect most genetic studies; in the present system, however, a misincorporated base would change a normal (digestion-sensitive) PCR product into a digestion-resistant product, which would then be preferentially enriched during subsequent PCR, leading to a false positive digestion-resistant band. (Citations omitted.)

(*Id.*)

As noted above, it was known that mutant-enriched PCR techniques could achieve a sensitivity of at least 1 mutant in 10^4 normal cells (Day Declaration ¶ 9). However, this reported detection sensitivity was questioned in the literature, where it was noted that "false-positive results are possible as the error rate of *Taq* polymerase or other thermostable enzymes is surpassed." (*Id.*). In fact, when a mutant-enriched PCR protocol designed to limit false positives was used, one researcher was unable to improve upon the sensitivity of the mutation-ligation assay alone, reported to be just one mutant in 200 normal cells (*Id.*). Other scientists noted specifically that PCR-PIREMA is susceptible of error as high as one error per 10^4 bases under standard PCR conditions, and further modified the method to reduce the potential for PCR errors (*see* discussion of Oshita, below) (*Id.*).

A skilled scientist would have recognized that it is no trivial matter to introduce an RE site during PCR amplification (Day Declaration ¶ 10). Depending on context sequence and the specific base alterations comprising the mutant allele in question, mismatch primer extension, as in the case of Jacobson, is required to introduce an RE site into a nucleotide if a suitable site is not preexisting. (*Id.*). Mismatch primer extension is more or less difficult depending on the base changes attempted (*Id.*). Mismatch primer extension can produce terrible artifacts, possibly resulting from primer slippage (*Id.*). Most artifacts of this variety will result in PCR products that are refractory to RE cleavage (*Id.*). Furthermore, the potential for false positive sequence generation is actually greater than described in Jacobson (*Id.*). Jacobson states that the expected rate of error production is one base in 8000 for *Taq* polymerase at best (*Id.*). What is not said, but understood by a skilled scientist, is that this error rate exists for each PCR cycle (*Id.*). Consequently, the errors accumulate for every cycle that the PCR is performed (*Id.*). Furthermore, if it is assumed that a four-base restriction site is selected to be used in the PCR-PIREMA, then alteration of any of these 4 bases would result in a sequence uncleavable by the selected restriction endonuclease (*Id.*). Thus, the true false positive rate rises to 1 base in 2000 per cycle (*Id.*). After 20 cycles, the error rate is 1 base in 100, or 1%. After 60 cycles, error products with base alterations produced by polymerase misincorporation will have accumulated to comprise at least 3% of the PCR product (*Id.*). This problem of error accumulation is reflected in the attempts of others to alter PCR $MgCl_2$ and nucleotide concentrations, presumably to minimize the polymerase error (*Id.*). When the mutant allele is present in a sample in low abundance, error rates of this magnitude are simply unacceptable (*Id.*). Thus, the PCR/RE

technique of Jacobson would have been considered of limited utility to anyone skilled in the art (*Id.*).

In addition, it would have been understood by a skilled scientist that any reported "error rate" is actually an average of the polymerase error over a long stretch of nucleotides (Day Declaration ¶ 11). The error rate at a particular base in a given context sequence can be significantly different from the average, often as much as 10-fold different (*Id.*). Therefore, it is of paramount importance to minimize the background errors in a PCR based assay, particularly in an assay designed for the detection of a low abundance sequence in a mixed sample (*Id.*). Otherwise, while the method may work well in particular situations, sometimes deceptively well, it will fail catastrophically in others, making the assay unreliable (*Id.*). However, at the time Jacobson was published, the only techniques known to help minimize error rate involved adjusting nucleotide and $MgCl_2$ concentrations and/or annealing and synthesis times (*Id.*). Thus, an extensive amount of testing was carried out in making the present invention to find a reliable method of creating PCR products with the desired RE site (*Id.*). This included the determination of PCR extension efficiencies and specificities, with and without and mismatch conversion, using analogs and all four natural bases, and the testing of various proofreading and non-proofreading polymerases under a variety of buffer conditions (Example 7, Table 1 at pg. 49; *also see generally* Example 17, and Table 4 at pg. 70) (*Id.*). Without this information, a method such as the RE step of Jacobson was understood by a skilled scientist to be unreliable, and would not have been considered transferable to another method for identifying one or more low abundance sequences differing by one or more single-base changes in a sample containing a plurality of target nucleotide sequences (*Id.*).

Other sources of false positives during PCR amplification include cross-contamination of reagents and PCR product carryover (Day Declaration ¶ 12). Carryover problems arise due to the fact that a single PCR cycle produces very large numbers of amplifiable molecules that can potentially contaminate subsequent amplifications of the same target sequence (*Id.*). It is particularly important to avoid carryover when the number of target molecules relative to other molecules in the sample is small, which is almost universally the case when detecting mutant molecules in a mixed DNA sample (*Id.*). Because PCR amplification is exponential, false positives are exponentially increased with every additional PCR step (*Id.*). A typical completed PCR has approximately 10^{13} molecules of

DNA per ml, making even an inadvertent transfer of a small volume of an unamplified sample quite serious (*Id.*). For example, even a 0.01 µl carryover of a sample with 10^{13} molecules of DNA per ml contains 10^{8} molecules of DNA; 10^{8} molecules of DNA carried over into a new PCR reaction and further amplified poses a significant risk to the validity of the PCR results (*Id.*). Thus, it would have been clear to a scientist that using the multiple PCR/RE steps taught by Jacobson introduced the potential for significant detection error, which required burdensome validation experiments, additional internal controls, and costly repetition to ensure the validity of the results (*Id.*). A skilled scientist would have understood that a reproducible signal does not guarantee a reliable result (*Id.*).

In Jacobson, the sensitivity of the assay appears to be improved when a gel electrophoresis step is added following two rounds of PCR/RE/digest, and before a final PCR/RE/digest is carried out (Day Declaration ¶ 13). However, gel electrophoresis and the subsequent gel purification steps are laborious and time-consuming (*Id.*). More importantly, they also provide numerous opportunities for introducing DNA contamination into the assay (*Id.*). Aerosolization of samples, well-well leaking, and carry-over from pipette tips and the razor blade(s) used to excise the gel bands are all potential sources of contamination (*Id.*). Because the gel purified sample is subsequently subjected to one or more additional PCR steps ("additional enrichment") in Jacobson, any DNA contamination resulting from the gel electrophoresis and purification steps will be subject to an exponential increase in the ensuing reactions (*Id.*). Therefore, although Jacobson indicates that the gel purification step increased sensitivity in some instances, a skilled scientist would have been aware that these steps, in and of themselves, add a significant risk of false positives (*Id.*). In other words, with or without the gel steps, false positives are a problem inherent in the method of Jacobson. Furthermore, the gel steps did not predictably increase assay sensitivity (*Id.*). This is most likely because 1) it is difficult to control the multiple sources of contamination in the gel electrophoresis and purification steps and 2) DNA contamination arising at the gel steps is exponentially increased in the subsequent PCR step(s), producing variations in detection sensitivity from one experiment to the next (*Id.*). This lack of reliability would have been recognized by a skilled scientist, and, coupled with the potential of Jacobson for generating false positives, would have led a scientist to view Jacobson as unsuitable for use in conjunction with a method for identifying one or more low abundance sequences differing by one or more single-base changes in a sample containing a plurality of target nucleotide sequences (*Id.*).

The myriad of pitfalls of Jacobson have not gone unnoticed by the scientific community (Day Declaration ¶ 14). The method of Jacobson, even as later modified in Mills et al., "Detection of K-*ras* Oncogene Mutations in Bronchoalveolar Lavage Fluid for Lung Cancer Diagnosis," *J. National Cancer Institute* 87(14):1056-1060 (1995) ("Mills II"), attached to the Day Declaration as Exhibit 10, was contemporaneously criticized as "suffer[ing] from certain statistical limitations and selection bias, which do not allow us to accurately ascertain the potential clinical impact of these assays." (*Id.*). In discussing the *ras* mutation studies of Mills et al., Birrer goes on to note that there were problems associated with this type of assay, including "technical obstacles that need to be overcome to develop an assay that is both sensitive and specific at detecting the marker of interest." Birrer, M.J., "Translational Research and Epithelial Carcinogenesis: Molecular Diagnostic Assays Now-Molecular Screening Assays Soon?" *J. National Cancer Institute* 87(14):1041-1043 (1995) at pg. 1041, para. bridging left and rt. cols. ("Birrer"), attached to Day Declaration as Exhibit 9). According to Birrer, as detection assays become increasingly sensitive, involving more cycles of PCR amplification and enrichment, the possibility of artifactual mutation in PCR-RFLP increases (Day Declaration ¶ 14). In addition, because the error rate of *Taq* polymerase is one error per 10^4 bases, *Taq*-induced mutations are a concern (*Id.*). As discussed above, polymerase mutations can arise due to misincorporation, leading to the amplification of false positives in the sample (*Id.*).

Furthermore, the fact that the Jacobson method of detection was problematical is made abundantly clear in two journal articles, of which Jacobson is a co-author, that followed the publication of Jacobson: Mills et al., "Increased Prevalence of K-*ras* Oncogene Mutations in Lung Adenocarcinoma," *Cancer Research* 55:1444-1447 (1995) ("Mills I") attached to the Day Declaration as Exhibit 6 and Mills II (Day Declaration ¶ 15). In Mills I, the PCR-PIREMA method of Jacobson for the detection of mutant *ras*-genes was followed, with some "important modifications" in the PCR steps, including the decrease in the nucleotide and $MgCl_2$ concentrations and lengthened annealing and synthesis times (*Id.*). In addition, the gel purification step was eliminated, which simplified the procedure but decreased the sensitivity of the assay (*Id.*). Mills II followed the modified procedure described in Mills I, taking, in addition, "extensive measures" to avoid cross-contamination of the PCR samples (*Id.*). All of these subsequent modifications to the method of Jacobson indicate that the method was so susceptible to error that the changes were required to make the assay practical and reliable (*Id.*). In fact, in the effort to produce a reliable assay, it

appears that Mills et al. abandoned the high-sensitivity version of the PCR-PIREMA protocol, given that the mutation detection level of the modified version of Jacobson taught by Mills I-II is just 0.1% (Mills I at pg. 1444, rt. col., 2nd full para.), i.e., 1000-fold less sensitive than originally reported (*Id.*). This level of detection was already available in a number of other methodologies in use at the time (*Id.*). Thus, neither the original method of Jacobson, fraught with the potential for false positives, nor the method of Jacobson as modified by Mills I-II, provided any apparent benefit over other known methods for identifying one or more low abundance sequences differing by one or more single-base changes in a sample containing a plurality of target nucleotide sequences (*Id.*).

Thus, it would have been clear to a skilled scientist that by adding multiple PCR/RE steps, along with one or more electrophoresis and gel purification steps as taught by Jacobson, a mutation detection sensitivity of 1 in 10^6 in a mixed DNA sample was *theoretically* possible (Day Declaration ¶ 16). However, the cost of this theoretical level of detection sensitivity was high: the additional steps were cumbersome, and more importantly, the method was subject to false positives due to contamination, carry over, and, in particular, polymerase error (*Id.*). This potential "cost" was too high to be suitable for inclusion in an assay in which a very small amount of error can destroy the reliability and accuracy of the procedure (*Id.*). Furthermore, the repeated PCR/RE steps and the need for gel purification make the method of Jacobson unsuitable for adaptation to a large-scale automated screening process (*Id.*). By eliminating the gel steps and modifying the PCR protocol as taught by Mills I-II, Jacobson could be streamlined, however, in doing so, the sensitivity of the assay was reduced to 1 in 10^3, and the method still required multiple PCR amplification steps and at least two RE digestions, presenting the same issue with regard to potential for polymerase error and a high rate of false positives (*Id.*).

Therefore, based on all the above, a skilled scientist would have concluded that adding an RE step to a PCR-based DNA mutation detection methodology would not provide any measurable benefit with regard to increased sensitivity over other DNA mutation detection methods known at the time, and furthermore, carried with it the very real potential for false positives (Day Declaration ¶ 17). In an assay in which a very small amount of error can destroy the reliability of the procedure, this would have been an unacceptable risk (*Id.*). Thus, there would have been no reason to use Jacobson's technique, modified or unmodified, in conjunction with a method for identifying one or more low abundance sequences differing

by one or more single-base changes in a sample containing a plurality of target nucleotide sequences (*Id.*).

Thus, for all the reasons described above, a skilled scientist would not have been motivated to combine Jacobson with Barany, let alone had a reasonable expectation of success in doing so.

Accordingly, the rejection of claims 1, 3-7, and 17 for obviousness over Barany in view of Jacobson is improper and should be withdrawn.

The rejection of claim 2 under 35 U.S.C. § 103(a) for obviousness over Barany in view of Jacobson and further in view of Day et al., "Detection of Steroid 21-Hydroxylase Alleles Using Gene-Specific PCR and a Multiplexed Ligation Detection Reaction," *Genomics* 29:152-162 (1995)("Day") is respectfully traversed.

Day discloses a method utilizing gene-specific PCR amplification in conjunction with thermostable DNA ligase to discriminate single nucleotide variations in a multiplexed ligation detection assay in which fluorescent or radioactive ligation products are detected by electrophoresis on denaturing acrylamide gels. Since Day does not overcome the deficiencies of Barany and Jacobson with regard to claim 1, the combination of Barany, Jacobson, and Day cannot form a proper basis for the rejection of claim 2.

Therefore, the rejection of claim 2 for obviousness over Barany in view of Jacobson and Day is improper and should be withdrawn.

The rejection of claims 8 and 10-16 under 35 U.S.C. § 103(a) for obviousness over Barany in view of Jacobson and U.S. Patent No. 5,859,221 to Cook et al. ("Cook") is respectfully traversed.

Cook discloses sugar-modified oligonucleotides and oligonucleotide analogs useful as therapeutics, diagnostics, and research agents. Cook does not teach or suggest using nucleotide analogs in "[a] method for identifying one or more low abundance sequences differing by one or more single-base changes, insertions, or deletions from a high abundance sequence, in a sample containing a plurality of target nucleotide sequences."

Because Cook does not overcome the deficiencies of Barany and Jacobson, the rejection of claims 8 and 10-16 over Barany in view of Jacobson and Cook is improper and should be withdrawn.

The rejection of claim 9 under 35 U.S.C. § 103(a) for obviousness over Barany in view of Jacobson, Cook, and Day is respectfully traversed, because Day does not overcome the above-noted deficiencies of Barany, Jacobson, and Cook. Therefore, this rejection of claim 9 is improper and should be withdrawn.

The rejection of claims 1, 2, 5-7, and 17 under 35 U.S.C. § 103(a) for obviousness over Belgrader et al., "A Multiplex PCR-Ligase Detection Reaction Assay for Human Identity Testing," *Genome Science Technology* 1(2):77-87 (1996) ("Belgrader") in view of Jacobson is respectfully traversed.

Belgrader teaches a coupled multiplex PCR-LDR assay to type single base variations at 12 biallelic loci, giving a power of discrimination of $1.1.2 \times 10^5$. Since Belgrader's assay is similar to Barany's, the rejection over Belgrader in view of Jacobson is improper for substantially the same reasons noted above for the rejection over Barany in view of Jacobson. Accordingly, the rejection of claims 1, 2, 5-7, and 17 over Belgrader in view of Jacobson is improper and should be withdrawn.

The rejection of claims 8, 9, and 12-16 under 35 U.S.C. § 103(a) for obviousness over Belgrader in view of Jacobson and Cook is respectfully traversed for substantially the reasons noted above. Therefore, this rejection of claims 8, 9, and 12-16 is improper and should be withdrawn.

In view of all of the foregoing, applicants submit that this case is in condition for allowance and such allowance is earnestly solicited.

Respectfully submitted,

Date: December 22, 2003

Michael L. Goldman
Registration No. 30,727

NIXON PEABODY LLP
Clinton Square, P.O. Box 31051
Rochester, New York 14603-1051
Telephone: (585) 263-1304
Facsimile: (585) 263-1600

Certificate of Mailing - 37 CFR 1.8(a)

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner for Patents P.O. Box 1450 Alexandria, VA 22313-1450, on the date below.

December 22, 2003 — Jo Ann Whalen
Date — Jo Ann Whalen

PATENT
Docket No.: 19603/481 (CRF D-2472A)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s) :	Barany et al.	Examiner: B. J. Forman
Serial No. :	09/528,014	Art Unit: 1634
Cnfrm. No. :	4478	
Filed :	March 17, 2000	
For :	COUPLED POLYMERASE CHAIN REACTION-RESTRICTION ENDONUCLEASE DIGESTION-LIGASE DETECTION REACTION PROCESS	

DECLARATION OF JOSEPH P. DAY UNDER 37 CFR § 1.132

Mail Stop RCE
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

I, JOSEPH P. DAY, hereby declare that:

1. I received a Ph.D. in Biophysics from the University of California, San Francisco, in 1993.

2. I am currently a Senior Bioinformatics Scientist at Applied Biosystems, Foster City, CA.

3. I am a co-inventor of the above-identified patent application.

4. This declaration is submitted to demonstrate that the mutant gene detection method described in Jacobson et al., "A Highly Sensitive Assay For Mutant *ras* Genes And Its Application To The Study of Presentation And Relapse Genotypes In Acute Leukemia," *Oncogene* 9:553-563 (1994) ("Jacobson"), would not have been viewed by a skilled scientist as a technique suitable for use in conjunction with a method for identifying one or more low abundance sequences differing by one or more single-base changes in a sample containing a plurality of target nucleotide sequences.

5. I am familiar with Jacobson, which teaches a method called PCR-Primer Introduced Restriction with Enrichment of Mutant Alleles ("PCR-PIREMA") for the detection of mutant *ras* genes. In PCR-PIREMA, the starting sample is amplified by the polymerase chain reaction ("PCR") technique using different "matched primers" to specifically amplify one of the known sites of mutations in a *ras* gene. A second PCR is carried out on the primary PCR extension product with mismatched "screening" primers to create a restriction endonuclease ("RE") site in the secondary extension product. The formation of the restriction site is dependent upon both introduced substitution and the normal codon sequence. Therefore, only normal alleles contain the introduced restriction site. The secondary extension products are digested with a restriction endonuclease recognizing the newly-created RE site, thereby reducing the presence of full-length normal sequences in the sample. Next, enrichment of the mutant allele population is carried out by subjecting the sample to a third PCR using the "screening primer" of the second PCR step, followed by a second restriction digestion, to "enrich" the sample for mutant alleles. The sample is subjected to gel electrophoresis, and the full-length (digestion-resistant) mutant bands are gel-purified. In an unspecified number of reactions, an "additional enrichment" step consisting of a third PCR/RE reaction and digest is carried out after the gel purification step. The digestion-resistant band indicates the presence of a mutant allele, but does not define the sequence of a mutant allele. To demonstrate that a mutant allele was present originally, the gel-purified band is further amplified and digested using the same "screening" primer and restriction enzyme as used in the secondary PCR step (i.e., "enriched screening). To define the mutation, yet another PCR is performed on the gel-purified product using mismatched "verification" primers that introduce new restriction sites into the PCR products derived from specific mutant alleles, the opposite of the screening strategy. The presence of a PCR product that cuts with the verification enzyme confirms that a specific mutation is present.

6. A scientist, considering the use of Jacobson's technique in conjunction with a method for identifying one or more low abundance sequences differing by one or more single-base changes in a sample containing a plurality of target nucleotide sequences, would need to know whether the use of that technique achieved an increase in detection sensitivity, and if so, whether such increased sensitivity had a detrimental effect. As I will describe in greater detail below, the method of Jacobson has a serious shortcoming, in particular, a high potential for false positive results, which a skilled scientist would have recognized.

Accordingly, there would be no reason to use Jacobson's technique in conjunction with a method for identifying one or more low abundance sequences differing by one or more single-base changes in a sample containing a plurality of target nucleotide sequences.

7. PCR-PIREMA is described by Jacobson as an "extremely sensitive" method of detecting mutant alleles in a biological sample having both normal and, potentially, mutant alleles of a particular DNA (Jacobson, pg. 555, left col., 2nd full para.). However, what a skilled scientist would have appreciated about Jacobson is that in order to reach a detection level of one mutant allele in 1,000,000 normal alleles (1 in 10^6) that Jacobson asserts PCR-PIREMA is capable of achieving, a single PCR-PIREMA sample, as described above, is subjected to PCR no fewer than three times, at least two RE digestion steps, one or more gel electrophoresis steps, and one or more gel purification procedures (Jacobson, para. bridging pp. 557-559). These multiple steps make the method of Jacobson unattractive for several reasons.

False Positives A Known Problem in Repetitive PCR

8. Firstly, as described in my application (pg. 7, line 9 to pg. 8, line 23) and by others, the method of Jacobson is an adaptation of the commonly used PCR-based gene analysis technique known as PCR-RFLP (analysis by generation of artificial restriction fragment length polymorphisms), which requires repetitive steps to ensure an enrichment for allelic variants (*see, e.g.*, Behn et al., "Facilitated Detection of Oncogene Mutations from Exfoliated Tissue Material By PNA-mediated 'Enriched PCR' Protocol," *J. Pathology* 190:69-75 (2000), pg. 69, 2nd full para., ("Behn"), attached hereto as Exhibit 1). Although PCR-RFLP-based protocols achieve a relatively high sensitivity for mutant alleles, up to 1 in 10^4, the literature teaches that this method bears an increasing risk of polymerase-borne infidelity as the error rate of *Taq* or other thermostable polymerases is surpassed following repeated cycling. (*Id.*, *see also* Ahrendt et al., "Molecular Detection of Tumor Cells in Bronchoalveolar Lavage Fluid from Patients with Early Stage Lung Cancer," *J. National Cancer Institute* 91(4):332-339, at pg. 337, 2nd full para. (1999) ("Ahrendt"), attached hereto as Exhibit 2). Errors due to infidelity have been found to be introduced early in the PCR process. Therefore, when the total number of PCR cycles involved exceeds approximately 70-80 cycles, as in Jacobson, the likelihood of error is significant (Jacobson at 559, left col., 1st full para.). The rate of misincorporation is not a problem in most PCR systems. However, when the concentration of the desired DNA is low in the starting sample, as is the case in the

PCR-PIREMA protocol of Jacobson, misincorporation is a serious problem, as acknowledged by Jacobson:

> In theory, another potential pitfall to this method is that after enrichment, a digestion resistant band could contain PCR products which had misincorporated a base during 1° or 2° PCR; such products might arise because *Taq* polymerase has a misincorporation rate estimated at 1/6000-1/8000 bases. This rate does not affect most genetic studies; in the present system, however, a misincorporated base would change a normal (digestion-sensitive) PCR product into a digestion-resistant product, which would then be preferentially enriched during subsequent PCR, leading to a false positive digestion-resistant band. (Citations omitted.)

Jacobson at 559, rt. col., 1st full para.

9. As noted above, it was known that mutant-enriched PCR techniques could achieve a detection sensitivity of at least 1 mutant in 10^4 normal cells (Ahrendt at pg. 337, 2nd full para., citing, among others, the K-*ras* mutation assay of Jacobson). However, this reported detection sensitivity was questioned by Ahrendt, which goes on to say that "false-positive results are possible as the error rate of *Taq* polymerase or other thermostable enzymes is surpassed." (*Id.*). In fact, when a mutant-enriched PCR protocol designed to limit false positives was used, Ahrendt was unable to improve upon the sensitivity of the mutation-ligation assay alone, reported to be just one mutant in 200 normal cells (*Id.*). Other scientists noted specifically that PCR-PIREMA is susceptible of error as high as one error per 10^4 bases under standard PCR conditions (Oshita et al., "Detection of K-ras Mutations of Bronchoalveolar Lavage Fluid Cells Aids the Diagnosis of Lung Cancer in Small Pulmonary Lesions, *Clinical Cancer Research* 5:617-620, (1999) at pg. 619, right col., 1st full para. ("Oshita"), attached hereto as Exhibit 3) and further modified the method to reduce the potential for PCR errors (*see* discussion of Oshita in ¶ 14, below).

10. A skilled scientist would have recognized that it is no trivial matter to introduce an RE site during PCR amplification. Depending on context sequence and the specific base alterations comprising the mutant allele in question, mismatch primer extension, as in the case of Jacobson, is required to introduce an RE site into a nucleotide if a suitable site is not preexisting. Mismatch primer extension is more or less difficult depending on the base changes attempted (Day et al., "Nucleotide Analogs Facilitate Base Conversion with 3'

Mismatch Primers," *Nucleic Acids Res.* 27(8):1810-8 (April 15, 1999) ("Day I"), attached hereto as Exhibit 4; and Day et al., "Nucleotide Analogs and New Buffers Improve a Generalized Method to Enrich for Low Abundance Mutations," *Nucleic Acids Res.* 27(8):1819-27 (April 15, 1999) ("Day II"), attached hereto as Exhibit 5). Mismatch primer extension can produce terrible artifacts, possibly resulting from primer slippage. Most artifacts of this variety will result in PCR products that are refractory to RE cleavage. Furthermore, the potential for false positive sequence generation is actually greater than described in Jacobson. Jacobson states that the expected rate of error production is one base in 8000 for *Taq* polymerase at best (Jacobson at 599, right col, 2nd full para.). What is not said, but understood by a skilled scientist, is that this error rate exists for each PCR cycle. Consequently, the errors accumulate for every cycle that the PCR is performed. Furthermore, if it is assumed that a four-base restriction site is selected to be used in the PCR-PIREMA, then alteration of any of these 4 bases would result in a sequence uncleavable by the selected restriction endonuclease. Thus, the true false positive rate rises to 1 base in 2000 per cycle. After 20 cycles, the error rate is 1 base in 100, or 1%. After 60 cycles, error products with base alterations produced by polymerase misincorporation will have accumulated to comprise at least 3% of the PCR product. This problem of error accumulation is reflected in the attempts of others to alter PCR $MgCl_2$ and nucleotide concentrations (Mills et al., "Increased Prevalence of K-*ras* Oncogene Mutations in Lung Adenocarcinoma," *Cancer Research* 55:1444-1447 (1995) ("Mills I") at pg 1444, rt. 2nd full para., attached hereto as Exhibit 6), presumably to minimize the polymerase error (*also see generally* Ling LL, et al., "Optimization of the Polymerase Chain Reaction With Regard to Fidelity: Modified T7, Taq, and Vent DNA Polymerases," *PCR Methods Appl.* 1(1):63-9 (1991) ("Ling"), attached hereto as Exhibit 7). When the mutant allele is present in a sample in low abundance, error rates of this magnitude are simply unacceptable. Thus, the PCR/RE technique of Jacobson would have been considered of limited utility to anyone skilled in the art.

11. In addition, it would have been understood by a skilled scientist that any reported "error rate" is actually an average of the polymerase error over a long stretch of nucleotides. The error rate at a particular base in a given context sequence can be significantly different from the average, often as much as 10-fold different. Therefore, it is of paramount importance to minimize the background errors in a PCR based assay, particularly in an assay designed for the detection of a low abundance sequence in a mixed sample. Otherwise, while the method may work well in particular situations, sometimes deceptively

well, it will fail catastrophically in others, making the assay unreliable. However, at the time Jacobson was published, the only techniques known to help minimize error rate involved adjusting nucleotide and $MgCl_2$ concentrations and/or annealing and synthesis times (*e.g.*, Mills I at pg. 1444, rt. col., 2nd full para.). Thus, an extensive amount of testing was carried out in the making of my invention to find a reliable method of creating PCR products with the desired RE site. This included the determination of PCR extension efficiencies and specificities, with and without and mismatch conversion, using analogs and all four natural bases, and the testing of various proofreading and non-proofreading polymerases under a variety of buffer conditions (Example 7, Table 1 at pg. 49; *also see generally* Example 17, and Table 4 at pg. 70). Without this information, a method such as the RE step of Jacobson was understood by a skilled scientist to be unreliable, and would not have been considered transferable to another method for identifying one or more low abundance sequences differing by one or more single-base changes in a sample containing a plurality of target nucleotide sequences.

12. Other sources of false positives during PCR amplification include cross-contamination of reagents and PCR product carryover. Carryover problems arise due to the fact that a single PCR cycle produces very large numbers of amplifiable molecules that can potentially contaminate subsequent amplifications of the same target sequence (Kwok et al., "Avoiding False Positives with PCR," Nature 339:237-238, correction in *Nature* 339:490 (1989), at 237, 2nd full para. ("Kwok"), attached hereto as Exhibit 8). It is particularly important to avoid carryover when the number of target molecules relative to other molecules in the sample is small (*Id.*), which is almost universally the case when detecting mutant molecules in a mixed DNA sample. Because PCR amplification is exponential, false positives are exponentially increased with every additional PCR step. A typical completed PCR has approximately 10^{13} molecules of DNA per ml, making even an inadvertent transfer of a small volume of an unamplified sample quite serious. For example, even a 0.01 µl carryover of a sample with 10^{13} molecules of DNA per ml contains 10^8 molecules of DNA; 10^8 molecules of DNA carried over into a new PCR reaction and further amplified poses a significant risk to the validity of the PCR results. Thus, it would have been clear to a scientist that using the multiple PCR/RE steps taught by Jacobson introduced the potential for significant detection error, which required burdensome validation experiments, additional internal controls, and costly repetition to ensure the validity of the results. A skilled scientist would have understood that a reproducible signal does not guarantee a reliable result.

13. In Jacobson, the sensitivity of the assay appears to be improved when a gel electrophoresis step is added following two rounds of PCR/RE/digest, and before a final PCR/RE/digest is carried out (Jacobson at pg. 559, lines 9-12). However, gel electrophoresis and the subsequent gel purification steps are laborious and time-consuming. More importantly, they also provide numerous opportunities for introducing DNA contamination into the assay. Aerosolization of samples, well-well leaking, and carry-over from pipette tips and the razor blade(s) used to excise the gel bands are all potential sources of contamination. Because the gel purified sample is subsequently subjected to one or more additional PCR steps ("additional enrichment") in Jacobson, any DNA contamination resulting from the gel electrophoresis and purification steps will be subject to an exponential increase in the ensuing reactions. Therefore, although Jacobson indicates that the gel purification step increased sensitivity in some instances (Jacobson at 559, left col., lines 9-11), a skilled scientist would have been aware that these steps, in and of themselves, add a significant risk of false positives. In other words, with or without the gel steps, false positives are a problem inherent in the method of Jacobson. Furthermore, the gel steps did not predictably increase assay sensitivity (*Id.*, describing improved sensitivity as occurring "often".) This is most likely because 1) it is difficult to control the multiple sources of contamination in the gel electrophoresis and purification steps and 2) DNA contamination arising at the gel steps is exponentially increased in the subsequent PCR step(s), producing variations in detection sensitivity from one experiment to the next. This lack of reliability would have been recognized by a skilled scientist, and, coupled with the potential of Jacobson for generating false positives, would have led a scientist to view Jacobson as unsuitable for use in conjunction with a method for identifying one or more low abundance sequences differing by one or more single-base changes in a sample containing a plurality of target nucleotide sequences.

14. The myriad of pitfalls of Jacobson have not gone unnoticed by the scientific community. The method of Jacobson, even as later modified in Mills II (see discussion below in ¶ 15), was contemporaneously criticized as "suffer[ing] from certain statistical limitations and selection bias, which do not allow us to accurately ascertain the potential clinical impact of these assays." Birrer, M.J., "Translational Research and Epithelial Carcinogenesis: Molecular Diagnostic Assays Now-Molecular Screening Assays Soon?" *J. National Cancer Institute* 87(14):1041-1043 (1995) at pg. 1041, para. bridging left and rt. cols. ("Birrer"), attached hereto as Exhibit 9). In discussing the *ras* mutation studies

of Mills et al., Birrer goes on to note that there were problems associated with this type of assay, including "technical obstacles that need to be overcome to develop an assay that is both sensitive and specific at detecting the marker of interest." (*Id.* at pg. 1041. rt. col., 1st full para.). According to Birrer, as detection assays become increasingly sensitive, involving more cycles of PCR amplification and enrichment, the possibility of artifactual mutation in PCR-RFLP increases (*Id.* at pg. 1042, 1st full para.). In addition, because the error rate of *Taq* polymerase is one error per 10^6 bases, *Taq*-induced mutations are a concern (*Id.*, *see also* Oshita at pg 619, right col., 1st full para.). As discussed above, polymerase mutations can arise due to misincorporation, leading to the amplification of false positives in the sample.

Modification of Jacobson Raises Issue of Usefulness and Reliability

15. Furthermore, the fact that the Jacobson method of detection was problematical is made abundantly clear in two journal articles, of which Jacobson is a co-author, that followed the publication of Jacobson: Mills I (Exhibit 6), and Mills et al., "Detection of K-*ras* Oncogene Mutations in Bronchoalveolar Lavage Fluid for Lung Cancer Diagnosis," *J. National Cancer Institute* 87(14):1056-1060 (1995) ("Mills II"), attached hereto as Exhibit 10). In Mills I, the PCR-PIREMA method of Jacobson for the detection of mutant *ras*-genes was followed, with some "important modifications" in the PCR steps, including the decrease in the nucleotide and $MgCl_2$ concentrations and lengthened annealing and synthesis times (Mills I at pg. 1444, rt. col., 2nd full para.). In addition, the gel purification step was eliminated, which simplified the procedure but decreased the sensitivity of the assay. Mills II followed the modified procedure described in Mills I, taking, in addition, "extensive measures" to avoid cross-contamination of the PCR samples (Mills II, pg. 1057, rt. col., 5th full para.). All of these subsequent modifications to the method of Jacobson indicate that the method was so susceptible to error that the changes were required to make the assay practical and reliable. In fact, in the effort to produce a reliable assay, it appears that Mills et al. abandoned the high-sensitivity version of the PCR-PIREMA protocol, given that the mutation detection level of the modified version of Jacobson taught by Mills I-II is just 0.1% (Mills I at pg. 1444, rt. col., 2nd full para.), i.e., 1000-fold less sensitive than originally reported. This level of detection was already available in a number of other methodologies in use at the time (*see e.g.*, Jacobson, at 557, last para., over to pg. 559, naming several references in which $1:10^3$ to $1:10^4$ detection sensitivity was obtained). Thus, neither the original method of Jacobson, fraught with the potential for false positives,

nor the method of Jacobson as modified by Mills I-II, provided any apparent benefit over other known methods method for identifying one or more low abundance sequences differing by one or more single-base changes in a sample containing a plurality of target nucleotide sequences.

16. Thus, it would have been clear to a skilled scientist that by adding multiple PCR/RE steps, along with one or more electrophoresis and gel purification steps as taught by Jacobson, a mutation detection sensitivity of 1 in 10^6 in a mixed DNA sample was *theoretically* possible (Jacobson at 557-559, bridging para.). However, the cost of this theoretical level of detection sensitivity was high: the additional steps were cumbersome, and more importantly, the method was subject to false positives due to contamination, carry over, and, in particular, polymerase error. This potential "cost" was too high to be suitable for inclusion in an assay in which a very small amount of error can destroy the reliability and accuracy of the procedure. Furthermore, the repeated PCR/RE steps and the need for gel purification make the method of Jacobson unsuitable for adaptation to a large-scale automated screening process. By eliminating the gel steps and modifying the PCR protocol as taught by Mills I-II, Jacobson could be streamlined, however, in doing so, the sensitivity of the assay was reduced to 1 in 10^3, and the method still required multiple PCR amplification steps and at least two RE digestions, presenting the same issue with regard to potential for polymerase error and a high rate of false positives.

17. Therefore, based on all the above, a skilled scientist would have concluded that adding an RE step to a PCR-based DNA mutation detection methodology would not provide any measurable benefit with regard to increased sensitivity over other DNA mutation detection methods known at the time, and furthermore, carried with it the very real potential for false positives. In an assay in which a very small amount of error can destroy the reliability of the procedure, this would have been an unacceptable risk. Thus, there would have been no reason to use Jacobson's technique, modified or unmodified, in conjunction with a method for identifying one or more low abundance sequences differing by one or more single-base changes in a sample containing a plurality of target nucleotide sequences.

18. I hereby declare that all statements made herein of my own knowledge are true; and further that these statements were made with the knowledge that willful false

statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that any such willful false statements may jeopardize the validity of the application or any patent issued thereon.

21 DEC, 2003

Joseph P. Day

EXHIBIT 1

Journal of Pathology
J Pathol 2000; 190: 69–75.

Original Paper

Facilitated detection of oncogene mutations from exfoliated tissue material by a PNA-mediated 'enriched PCR' protocol

Moira Behn[1], Christian Thiede[2], Andreas Neubauer[2], Wulf Pankow[3] and Marcus Schuermann[1]*

[1] Zentrum Innere Medizin, Philipps-Universität Marburg, Baldingerstrasse, D-35033 Marburg, Germany
[2] Medizinische Klinik I, Technische Universität Dresden, Fetscherstrasse 74, D-01307 Dresden, Germany
[3] Abt. Innnere Medizin III, Schwerpunkt Pneumonologie/Infektiologie, Krankenhaus Neukölln, Rudower Strasse 48, D-12313 Berlin, Germany

*Correspondence to: Marcus Schuermann, Zentrum Innere Medizin, Philipps-Universität Marburg, Baldingerstrasse, D-35033 Marburg, Germany.
E-mail: schuermann@mailer.uni-marburg.de

Received: 12 October 1998
Revised: 21 April 1999
Accepted: 1 September 1999

Abstract

An 'enriched polymerase chain reaction (PCR)' protocol has been established for the sensitive detection of oncogene mutations in body fluid samples from cancer patients. This two-step protocol combines an allele-specific PCR clamping step followed by a PCR-RFLP (restriction fragment length polymorphism) confirmatory step. The method thus resembles a nested PCR technique starting directly from genomic DNA material and, in no more than 54 PCR cycles, allows the sensitive detection of one mutant allele in 10^5 normal alleles. This protocol was tested on bronchial cytology samples and sputum taken from lung cancer patients and point mutations could be detected both in codon 12 of K-*ras* and in three codons (248, 249, and 273) of the *p53* gene. Comparing this protocol with a different 'enriched PCR' method based on repetitive PCR-RFLP steps, a high concordance was noted between the two methods. Although the present protocol seems to be less sensitive by approximately one order of magnitude, it is much easier to perform and thus could be applied to the rapid but sensitive detection of allelic subfractions in a population of cells derived from exfoliative material. Copyright © 2000 John Wiley & Sons, Ltd.

Keywords: PCR clamping; PNA; PCR-RFLP; 'enriched PCR'; *ras*; *p53*; lung cancer

Introduction

Activating point mutations of the *ras* and *p53* type can be used as genetic markers to screen for the presence of neoplastic cells in a number of different body fluids [1,2]. This has been largely facilitated by the development of 'enriched' PCR methods, which are able to distinguish between wild-type and mutant alleles and which also provide a high degree of sensitivity to detect a few mutant alleles in a large excess of normal DNA [3,4]. The high frequency of activating K-*ras* and *p53* mutations found in most solid tumours, and the fact that a large percentage of these are found within a relatively small number of so-called 'hot spot' positions, has encouraged the development of PCR techniques which focus on these areas of interest and at the same time provide a high degree of sensitivity [5].

To date, two PCR-based techniques have been most commonly used: allele-specific PCR, also referred to as PCR-ASA (allele-specific amplification), which directly aims at the selective amplification of mutant alleles [6], and PCR-RFLP (analysis by generation of artificial restriction fragment length polymorphisms), which requires repetitive steps to ensure an enrichment for allelic variants [4,7–9]. PCR-ASA has been shown to provide reasonable sensitivity for mutant sequences in the range of 1–0.01% of the total DNA, but is restricted to the analysis of individual codon-specific mutations and requires several allele-specific oligonucleotides in the analysis of each codon to cover all mutations in a given position [10]. PCR-RFLP avoids the necessity for multiple allele-specific primers by working with mismatch primers which generate a palindromic sequence around the codon to be analysed in all wild-type alleles, which can then be cut by the appropriate restriction enzyme. Alleles with mutations in any of the first two triplet positions of the codon examined, however, stay digestion-resistant and can thus be separated in a subsequent agarose gel electrophoresis. To allow a mutation-selective enrichment, the digested PCR product is then subjected to repetitive rounds of PCR-RFLP, thus yielding a high sensitivity for mutant alleles up to >0.01% of the total number of amplified DNA alleles [4]. Although PCR-RFLP-based protocols ensure a relatively high sensitivity, this method bears an increasing risk of *Taq* polymerase-borne infidelity as soon as the total number of PCR cycles involved exceeds approximately 70–80 cycles, depending on the *Taq* enzyme used and the presence of additional 'proof-reading' (3' to 5' exonuclease) activity in the assay [11]. Due to the number of repetitive PCR steps, RFLP-based 'enriched' PCR is time-consuming and has not therefore been applicable for routine clinical analysis.

More recently, the mutation-sensitive hybridization profile of peptide nucleic acids (PNAs) has been exploited to design protocols that allow the suppres-

ccc 0022-3417/2000/010069–07$17.50

sion of normal genomic sequences during amplification (referred to as PCR clamping) [12–14]. PNAs are oligonucleotide mimics containing a peptide instead of a ribose-phosphate backbone, which bind strongly and sequence specifically to complementary DNA strands by Watson–Crick hydrogen bonding [15–17]. In PCR clamping, a PNA oligomer binds to a target wild-type sequence and due to its inability to serve as a primer, blocks amplification of a DNA fragment confined by a pair of two externally located DNA oligonucleotide primers. In the case of a single-base mismatch, the DNA/PNA duplex is significantly destabilized, which allows strand elongation from a bound DNA oligomer to proceed, resulting in the detection of PCR fragments, most of which harbour the variant allele. PNA-directed PCR clamping can thus be a useful alternative for selective allele enrichment. A difficulty of this technique, however, is the fact that the allelic status of resulting PCR products cannot be distinguished except by time-consuming methods such as the sequencing of purified DNA fragments.

In the present study, we therefore took advantage of the allele-specific discrimination of PNAs, which we combined with a second (diagnostic) PCR step according to the PCR-RFLP technique. This method was termed CASE-PCR (combined allele-specific enriched PCR), according to the different character of each PCR step performed and the additive mutant selective effect achieved by this method. We applied this protocol to the specific enrichment of fragments mutated in several 'hot spot' regions of either the K-*ras* or the *p53* gene, which are frequently altered by codon-specific mutations in different types of human cancer. Our protocol ensures a comparatively high sensitivity, but also provides reliable detection of wild-type and mutant fragments in a single two-step reaction of no more than 54 PCR cycles, starting from genomic material.

Materials and methods

Cell lines, tissue source, and DNA preparation

DNA from cell lines NCI-H1573 and NCI-H2009 harbouring either a homozygous *p53* codon 248 mutation (NCI-H1573) or a heterozygous K-*ras* codon 12 and *p53* codon 273 mutation (NCI-H2009) was extracted as previously described [11]. Cytology, bronchial lavage, and sputum samples had been collected during routine bronchoscopy at the Neukölln Hospital Berlin and, for the analysis, had been treated as previously described [11].

DNA and PNA oligonucleotides

Oligonucleotide primers used in the preamplification and detection step were identical to those described by Behn *et al.* [11] PNA oligonucleotides were obtained from TIB MolBiol (Berlin, Germany) and used as described by Thiede *et al.* [13] and Behn and Schuermann [18]. In each case, the nucleotide sequence was identical to the non-coding strand of the wild-type allele and was chosen so that a 15-nucleotide segment around the 'hot spot' mutation site was covered:

PNAK12-3: H_2N-TACGCCACCAGCTCC-CON_2H
PNA248-9: H_2N-GGGCCTCCGGTTCAT-CON_2H
PNA273: H_2N-ACAAACACGCACCTC-CON_2H

PNA-PCR parameters (preamplification)

A 266 bp product spanning K-*ras* exon 1 was amplified using primers K-ras 5' and K-ras 3' and PNAK12-3 (illustrated schematically in Figure 1). For *p53* mutation analysis, a 647 bp DNA fragment containing exon 7 of the genomic *p53* was amplified using the primers p5 and p8 with PNA248-9 and a 350 bp fragment containing exon 8 was amplified using the primers p7 and p2 with PNA273 (see Figure 1). The PCR clamping reaction was performed according to Thiede *et al.* [13] and tested under conditions described previously [18,19]. Each reaction was performed in a volume of 25 µl, consisting of buffer [25 mM TAPS (pH 9.3), 50 mM KCl, 1.5 mM $MgCl_2$ and 1 mM DTT], 0.2 mM of each dNTP, 50 ng of DNA, 25 pmol of each primer, 75 pmol of PNA, 3.75% glycerol (7.5% glycerol in the case of PNAK12-3), and 1 unit of ELONGase® (Gibco BRL). To prevent non-specific polymerization prior to thermal cycling, hot start was performed by 5 min preheating at 94°C. The PCR itself was performed over 24 cycles with the following conditions: 60 s at 94°C; 60 s at 73°C (PNA248-9) or, alternatively, at 68°C (PNAs K12-3 and 273); 60 s at 54°C; and 120 s at 72°C. The additional 68°C/73°C step was chosen in order to allow preferential annealing of the PNAs to DNA.

Purification of PCR products

DNA fragments resulting from the first amplification were purified by centrifugation through quick spin columns (QIAquick PCR Purification Kit; Qiagen). The samples were then diluted 1 : 100 to 1 : 300 and 2 µl was taken for the PCR-RFLP analysis.

PCR-RFLP analysis

This step followed a previously described protocol [11]. Two microlitres of diluted preamplified products of the first step was amplified over 30 cycles in a volume of 50 µl containing the following buffer components: 60 mM Tris–SO_4 (pH 9.1), 18 mM $(NH_4)_2SO_4$, 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 25 pmol of each primer and, 2 Units of *Taq* DNA polymerase (Boehringer Mannheim). The following primer combinations were chosen (sequences described in ref. 11): K12Mva and K12as-1 for the detection of K-*ras* codon 12 mutations; primers 248CspI and as24x or 249Bsu36I and as24x to detect *p53* codon 248 and codon 249 mutations, respectively; and primers 273MluI and p8 for the detection of mutations in *p53* codon 273.

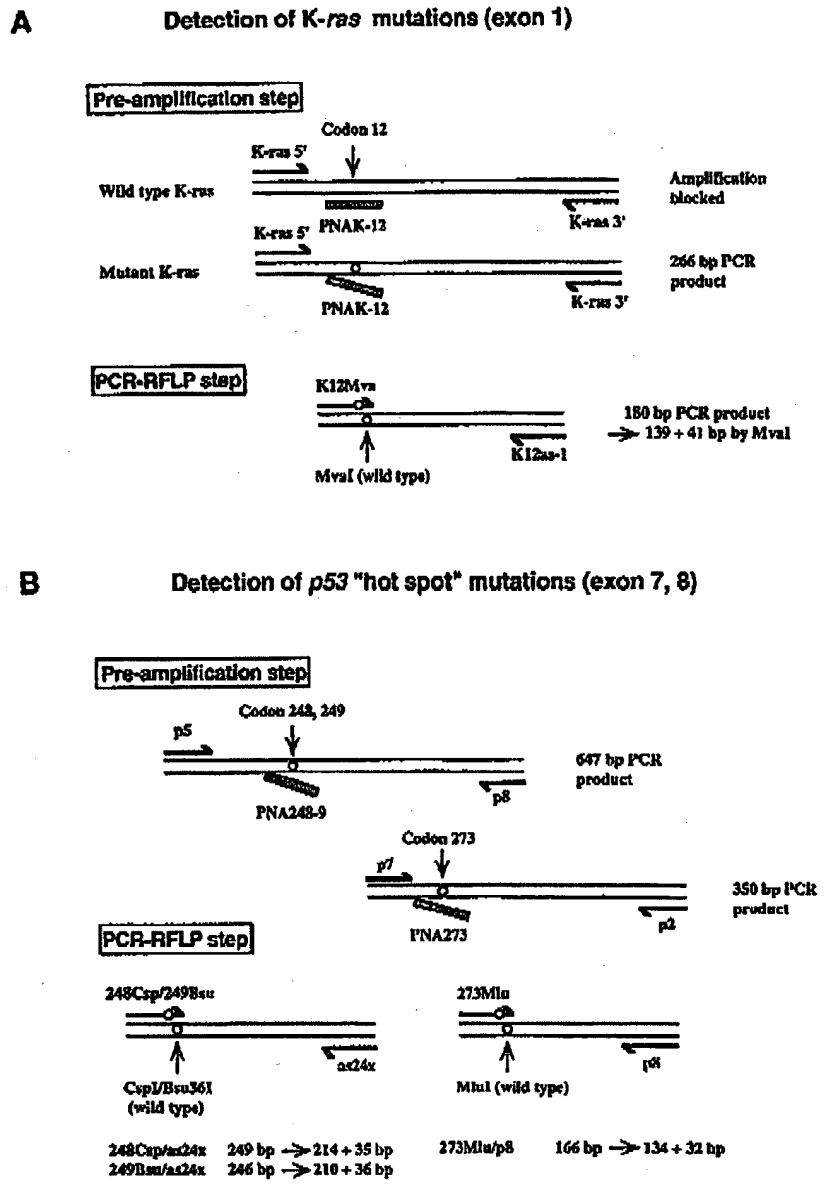


Figure 1. Schematic representation of the primers and PNAs used in the analysis. (A) Location of primers and PNAs used in the analysis of K-*ras*; (B) primers and PNAs used in the detection of *p53* lesions. The exon–intron structure of the genes is shown (not to scale), with exons given in boxes. Arrows indicate primer positions; the name of the primer is indicated above or below. The position of the PNAs is indicated by a bar. The primer pairs used in the PCR clamping step (primers and PNAs for first PCR) and the subsequent PCR–RFLP step (primers for PCR-RFLP) are listed below, along with the size of the expected and of the digested PCR products

Endonuclease digestion of PCR-RFLP products

Five-microlitre aliquots of the PCR-RFLP reaction were digested with 25 Units of the following respective enzymes. These were for K-*ras* codon 12: *Mva*I (Bochringer Mannheim); for *p53* codon 248: *Csp*I (Stratagene); for *p53* codon 249: *Bsu*36I (Stratagene); and for *p53* codon 273: *Mlu*I (Boehringer Mannheim). The digestion was performed for 3 h in a total volume of 25 µl under conditions recommended by the supplier (PCR reaction components were ignored); 20 µl of the digestion products was electrophoresed through a 3% ethidium bromide stained Nu Sieve® Agarose gel (Biozym, FMC Rockland).

Sequence analysis

For sequence analysis, PCR products were purified from agarose gels and sequenced according to standard 'cycle sequencing' protocols using an ABI PRISM 377 sequencing automate (Applied Biosystems), Ampli Taq DNA polymerase, and an ABI Prism Dye Terminator Cycle Sequencing Ready Reaction Kit according to the manufacturer's protocol.

Results and discussion

The enriched PCR protocol designed comprises two PCR steps, a first step involving a mutation-selective PNA oligonucleotide and conventional DNA primers intended to preamplify genomic material containing variant alleles of interest (shown schematically in Figure 1). The mutant-enriched PCR products are then reamplified using nested primers, one of which covers the site of mutation and generates a polymorphism recognized by a restriction endonuclease of choice (diagnostic PCR-RFLP step). Subsequent digestion will then allow the separation of all wild-type fragments by gel electrophoresis.

To test for the feasibility of the CASE-PCR protocol, we analysed DNA containing naturally occurring K-*ras* or *p53* mutations in distinct 'hot spot' positions. To allow comparison with previously established PCR-RFLP protocols, all exon-specific DNA primers for the preamplification of genomic material and the mismatch primers for the verification of K-*ras* mutations in codon 12 and *p53* mutations in codons 248, 249, and 273 were taken as previously described [11]. PNA concentrations and clamping conditions were initially tested for each individual oligomer and optimized as described in the Materials and methods section. We observed annealing of PNAs and specific clamping in a range between 70 and 76°C for PNA 248-9 and 64 and 70°C for PNAs K12-3 and 273. We determined the optimal PNA concentration to be around 38 μM (equivalent to 75 pmol per reaction).

With a limitation of 24–30 amplification cycles, the first step revealed specific PCR products of varying intensities, most of which were barely visible in ethidium bromide-stained gels (Figure 2, top panel). The second PCR-RFLP step then yielded PCR products of comparable intensities which were either completely digested, in the case of wild-type status, or digestion-resistant, in the case of a given K-*ras* or *p53* 'hot spot' mutation (Figure 2, bottom panel). The sensitivity of the combined PCR protocol was then tested by diluting DNA from a cell line harbouring a K-*ras* codon 12 or *p53* codon 248 mutation into DNA from a cell line with wild-type status. Figure 3 shows that the two-step protocol allowed the detection of up to one mutant allele in 10^3 wild-type alleles. Using a PCR-RFLP-based protocol consisting of four steps and a total of 102 PCR cycles, the same DNA dilution experiment revealed a sensitivity of approximately 1 allele in 10^4 (Figure 3, lower panel). Thus, the combined enriched PCR was approximately ten times less sensitive, although samples with a content of up to 0.1% mutant K-*ras* alleles could still be clearly distinguished from DNA containing only wild-type alleles (Figure 3). An analogous experiment in the absence of PNA probes revealed a limit of only 5% mutant K-*ras* alleles (Figure 2). We therefore estimate the specific enrichment effect of the PCR clamping step to be about 50-fold. Comparable sensitivities were obtained when PNAs 248-9 and 273 were used, designed to suppress the amplification of *p53* fragments with wild-type status around these positions.

We next analysed archival brush cytology material taken from 20 patients with bronchoscopically apparent lung cancer, which had been examined previously with respect to mutations occurring in either K-*ras* codon 12 or different *p53* 'hot spot' positions [11]. The results are shown in Figure 4. As the quality of DNA

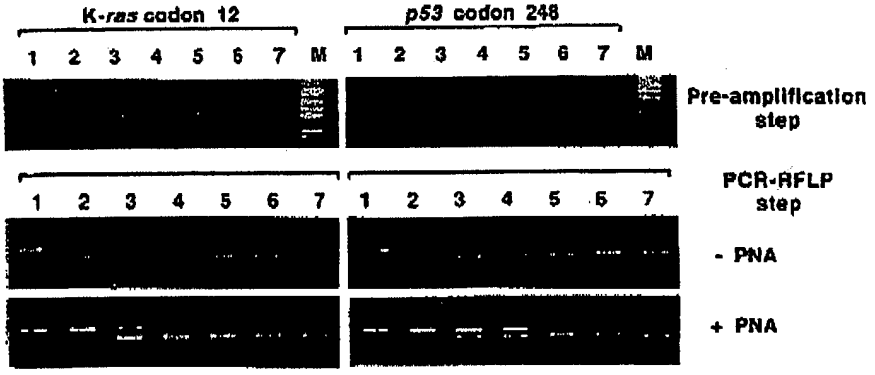


Figure 2. Allele-specific PNA clamping and subsequent detection by RFLP analysis. DNA from a cell line harbouring a homozygous K-*ras* codon 12 or *p53* codon 248 mutation was diluted into DNA from a cell line with wild-type K-*ras* (*p53*) status to yield 200 ng of genomic DNA per vial. Top: example of reaction products after 24 cycles of PNA clamping. Bottom: reaction products after PCR-RFLP (total of 48 cycles) either in the presence or in the absence of specific PNA. Lane 1, mutant cell line, no dilution; lane 2, diluted 1 : 10; lane 3, $1 : 10^2$; lane 4, $1 : 5 \times 10^2$; lane 5, $1 : 10^3$; lane 6, 5×10^2; lane 7, negative control

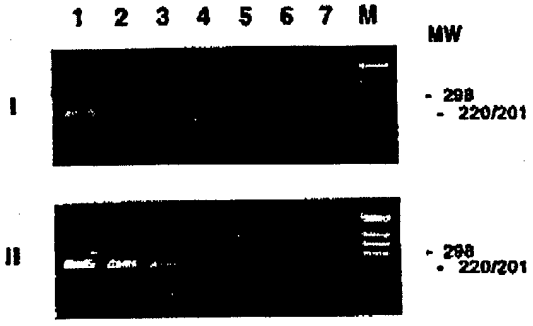


Figure 3. Serial dilution experiment comparing the sensitivity of CASE-PCR (top) and enrichment by repetitive PCR-RFLP (bottom). DNA from a cell line harbouring a homozygous K-*ras* codon 12 mutation was diluted into DNA from a cell line with wild-type K-*ras* status at the following concentrations: lane 1, no dilution; lane 2, diluted $1:10^2$; lane 3, $1:10^3$; lane 4, $1:3\times10^3$; lane 5, $1:5\times10^3$; lane 6, $1:9\times10^3$; lane 7, control without addition. CASE-PCR involved 54 cycles; 'mutant enriched' PCR-RFLP a total of 102 amplification cycles. To increase the proof-reading activity in both cases, the starting PCR was performed with ELONGase™, a mix which consists of *Taq* and *Pyrococcus* species GB-D thermostable DNA polymerases

was known to vary in these samples compared with DNA derived from cell lines, we obtained different amounts of final product. In all cases, however, we noticed a preponderance of mutant allele amplification in relation to wild-type fragments, showing the suppressive character of added wild-type complementary PNAs in the reaction. While nearly all lanes reveal digested PCR fragments indicative of wild-type allelostatus, a limited number of digestion-resistant fragments are clearly separated. In all cases, a previous analysis of the PCR products before restriction digestion (data not shown) and a respective control (DNA of a mutation-harbouring cell line diluted $1:10^3$) had been performed in parallel to ensure that amplification and digestion steps were not technically limiting.

The PNA-based mutant allele enrichment allowed us to detect five mutations in the K-*ras* gene at codon 12, while eight mutations were found in the *p53* gene at codons 248 (2), 249 (3), and 273 (3). Thus, a total of 13 mutations were detectable in cytology samples derived from eight patients (summarized in Table 1). When compared with the results obtained by an 'enriched' PCR-RFLP protocol previously applied, we found that the majority of lesions could be confirmed by our method (Table 1). Particularly when sequencing the digestion-resistant fragments, we found that all mutations detected by allele-specific clamping were also identical to those which had been found by the enriched PCR-RFLP approach. This result therefore underscores the reliability of our combined PCR protocol. A disadvantage was seen only with respect to the sensitivity of the method. We found only 13 mutations by CASE-PCR, while 19 lesions had been detected using repetitive PCR-RFLP-based enrichment. This can likely be attributed to the lower sensitivity of the current protocol, as most of the six mutations missed had become detectable only after three rounds of PCR-RFLP.

Finally, sputum samples derived from 23 lung cancer patients were examined using the combined enriched PCR method (summarized in Table 2). This time, we were able to detect five K-*ras* codon 12 mutations and three *p53* missense mutations occurring in codons 248, 249, and 273. Of eight lesions detected in this way, seven could be confirmed by analysis of lavage samples taken during bronchoscopy. As can be deduced from Table 2, in seven of eight cases the site of detection correlated with the site of tumour location, suggesting

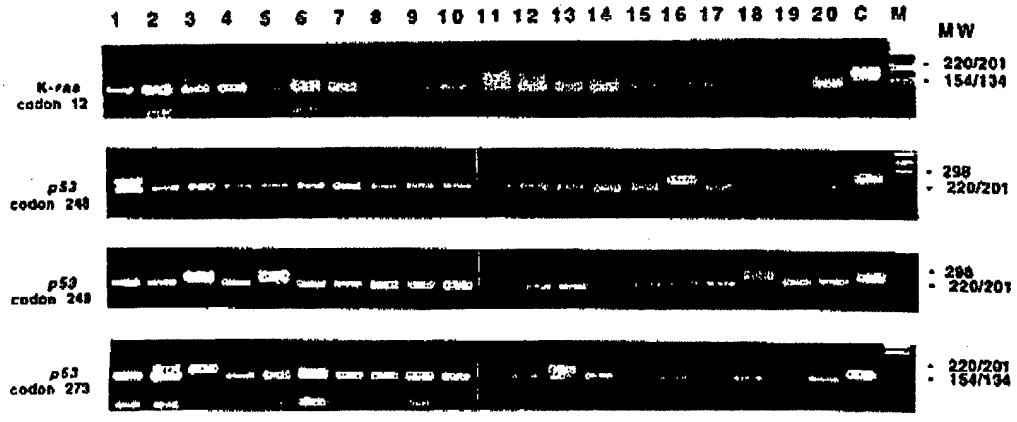


Figure 4. Analysis of 20 cytology samples for DNA containing K-*ras* or *p53* 'hot spot' mutations. CASE-PCR was performed as illustrated in Figure 1, starting with a 24-cycle PCR clamping reaction and subsequent nested PCR amplification over 30 cycles. Subsequent gel electrophoresis clearly separates residual wild-type alleles from enriched restriction-resistant mutant forms (running as retarded bands). C = DNA derived from a cell line containing the respective K-*ras* or *p53* point mutation in at least one allele; M = molecular DNA marker

Table 1. Detection of K-*ras* and *p53* mutations in brush cytology specimens by CASE-PCR and 'mutant enriched' PCR-RFLP

	CASE-PCR					'Mutant enriched' PCR-RFLP				
		p53					*p53*			
Sample No.	K-*ras* codon 12	codon 248	codon 249	codon 273		K-*ras* codon 12	codon 248	codon 249	codon 273	
1		CGG→CTG					CGG→CTG		CGT→CAT	
2				CGT→CTT			CGG→CAG		CGT→CTT	
3	GGT→GTT		AGG→AGC	CGT→CAT		GGT→GTT		AGG→AGC	CGT→CAT	
4										
5			AGG→AGC					AGG→AGC		
6										
7										
8	GGT→GTT					GGT→GTT	CGG→CTG			
9										
10										
11	GGT→GTT					GGT→GTT		AGG→AGC		
12	GGT→GCT					GGT→GCT				
13				CGT→CTT					CGT→CTT	
14							CGG→CTG			
15										
16		CGG→TGG					CGG→TGG			
17	GGT→GTT					GGT→GTT		AGG→AGC		
18			AGG→AGC					AGG→AGC		
19										
20										
Total	5	2	3	3	13	5	5	5	4	19

Table 2. Consistent detection of K-*ras* and *p53* mutations in lavage and sputum samples from lung cancer patients

			Lavage	
Patient No.	Tumour localization	Sputum	Right side	Left side
1	RUL	p53, 248		K-ras, 12
2	RUL			
3	RMB	K-ras, 12	K-ras, 12	
4	LMB			
5	LLL			
6	LMB			
7	RML	K-ras, 12	K-ras, 12	
8	LUL	p53, 249	p53, 249	p53, 249
9	RUL	K-ras, 12	K-ras, 12	
10	LUL			
11	LUL	K-ras, 12		K-ras, 12
12	RUL			
13	Multilocular			
14	LMB			
15	LUL			
16	LUL			
17	RUL	p53, 273	p53, 273	p53, 273
18	LMB			
19	RMB			
20	RMB			
21	RMB			
22	RUL	K-ras, 12		K-ras, 12
23	LMB			

RMB = right main bronchus; RUL = right upper lobe; RML = right middle lobe; LMB = left main bronchus; LUL = left upper lobe; LLL = left lower lobe.

that the mutations arose within the tumour cell mass. By sequence analysis we could also confirm that all digestion-resistant fragments harboured missense mutations in the respective codon (data not shown). Although sputum is thought to contain only a fraction of cells derived from the respiratory epithelium, which are at the same time of lower cytological quality than material obtained by lavage or bronchial brushing, the material is still sufficient to allow tumour-related DNA analysis. In this respect, we suggest that the CASE-PCR protocol might even provide a sensitive and additional method to look for distinct oncogene lesions in sputum obtainable by non-invasive methods.

From the results obtained so far, we conclude that CASE-PCR is a technically reliable method to screen for single codon oncogene mutations, with a sufficiently high sensitivity to account for low abundance alleles. Since it comprises only two PCR steps resembling a 'nested PCR', the protocol is much more rapid than conventional 'enriched' PCR-RFLP. The limitation to two successive PCR steps (48–54 PCR cycles) significantly reduces the risk of *Taq*-polymerase borne errors and at the same time allows a larger panel of samples to be processed for mutational screening.

Recently, a similar technique was described based on the combination of PNA clamping and allele-specific PCR [10]. While this method is even more sensitive, the ASA step requires a larger collection of allele-specific primers to cover all possible mismatches in each individual codon to be examined, which limits the

number of 'hot spot' positions that can be analysed. Moreover, different sequence-specific primer-template combinations occur in each case which need subtle adjustment of PCR conditions. The use of PNAs in the reaction generally offers the advantage to examine allelic mutations in several hot-spot positions (e.g. K-*ras*, codons 12 and 13; *p53*, codons 248 and 249), as long as these are covered by one PNA molecule. The combination with a second different PCR step in both protocols substantially improves the diagnostic impact since non-specific false-positive PCR products are avoided, similar to the benefit of 'nested PCR' protocols. In addition, we took care to allow DNA amplification to proceed over 24–30 PCR cycles, such as to ensure the gradual accumulation of a faint but specific PCR wild-type product in each case, despite the presence of 75 pmol of PNA in the reaction. This product could thus serve as an internal standard for efficient DNA amplification (see Figures 2 and 3). As stated earlier [11], the type of restriction enzyme chosen and the excessive amounts added in the reaction ensured a complete digest within 3 h. A disadvantage in our view, however, is the relatively small difference in size of digested and non-digested DNA fragments, which despite the use of high-resolution agarose gels (Nu Sieve® Agarose, Figure 4) makes it difficult to distinguish clearly both types of alleles. The use of polyacrylamide gels in this respect, combined with semi-automatic silver stain detection as applied more recently, did improve the discrimination of resulting DNA fragments but also amplified the signal of residual non-digested material due to improved sensitivity (data not shown). The overall benefit of this technique therefore remains questionable. While potential improvements remain to be tested, we nevertheless think that CASE-PCR in its present form already offers a practicable approach to allow more rapid diagnostic screening of a broader range of samples derived from exfoliative tissue material.

Acknowledgements

This work was supported by a grant donated by the Stiftung P.E. Kemkes and Stiftung VERUM to M.S.

References

1. Huebner K, Ohta M, Lubinski J, Berd D, Maguire HC Jr. Detection of specific genetic alterations in cancer cells. *Semin Oncol* 1996; 23: 22–30.
2. Loda M. Polymerase chain reaction-based methods for the detection of mutations in oncogenes and tumor suppressor genes. *Hum Pathol* 1994; 25: 564–571.
3. Mao L, Hruban RH, Boyle JO, Tockman M, Sidransky D. Detection of oncogene mutations in sputum precedes diagnosis of lung cancer. *Cancer Res* 1994; 54: 1634–1637.
4. Jacobson DR, Mills NE. A highly sensitive assay for mutant ras genes and its application to the study of presentation and relapse genotypes in acute leukemia. *Oncogene* 1994; 9: 553–563.
5. Greenblatt MS, Bennett WP, Hollstein M, Harris CC. Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. *Cancer Res* 1994; 54: 4855–4878.
6. Weber JL. Human DNA polymorphisms and methods of analysis. *Curr Opin Biotechnol* 1990; 1: 166–171.
7. Kahn SM, Jiang W, Culbertson TA, *et al*. Rapid and sensitive nonradioactive detection of mutant K-ras genes via 'enriched' PCR amplification. *Oncogene* 1991; 6: 1079–1083.
8. Levi S, Urbano-Ispizua A, Gill R, *et al*. Multiple K-ras codon 12 mutations in cholangiocarcinomas demonstrated with a sensitive polymerase chain reaction technique. *Cancer Res* 1991; 51: 3497–3502.
9. Mills NE, Fishman CL, Scholes J, Anderson SE, Rom WN, Jacobson DR. Detection of K-ras oncogene mutations in bronchoalveolar lavage fluid for lung cancer diagnosis. *J Natl Cancer Inst* 1995; 87: 1056–1060.
10. Rhodes CH, Honsinger C, Porter DM, Sorenson GD. Analysis of the allele-specific PCR method for the detection of neoplastic disease. *Diagn Mol Pathol* 1997; 6: 49–57.
11. Behn M, Qun S, Pankow W, Havemann K, Schuermann M. Frequent detection of ras and p53 mutations in brush cytology samples from lung cancer patients by a restriction fragment length polymorphism-based 'enriched PCR' technique. *Clin Cancer Res* 1998; 4: 361–371.
12. Ørum H, Nielsen PE, Egholm M, Berg RH, Buchardt O, Stanley C. Single base pair mutation analysis by PNA directed PCR clamping. *Nucleic Acids Res* 1993; 21: 5332–5336.
13. Thiede C, Bayerdorffer E, Blaszyk R, Wittig B, Neubauer A. Simple and sensitive detection of mutations in the ras proto-oncogenes using PNA-mediated PCR clamping. *Nucleic Acids Res* 1996; 24: 983–984.
14. Mrozikiewicz PM, Landt O, Cascorbi I, Roots I. Peptide nucleic acid-mediated polymerase chain reaction clamping allows allelic allocation of CYP1A1 mutations. *Anal Biochem* 1997; 250: 256–257.
15. Buchardt O, Egholm M, Berg RH, Nielsen PE. Peptide nucleic acids and their potential applications in biotechnology. *Trends Biotechnol* 1993; 11: 384–386.
16. Nielsen PE. DNA analogues with nonphosphodiester backbones. *Annu Rev Biophys Biomol Struct* 1995; 24: 167–183.
17. Agrawal S, Iyer RP. Modified oligonucleotides as therapeutic and diagnostic agents. *Curr Opin Biotechnol* 1995; 6: 12–19.
18. Behn M, Schuermann M. Sensitive detection of p53 gene mutations by a 'mutant enriched' PCR-SSCP technique. *Nucleic Acids Res* 1998; 25: 1356–1358.
19. Behn M, Schuermann M. Simple and reliable Factor V genotyping by PNA-mediated PCR-clamping. *Thromb Haemost* 1998; 79: 773–777.

EXHIBIT 2

ARTICLES

Molecular Detection of Tumor Cells in Bronchoalveolar Lavage Fluid From Patients With Early Stage Lung Cancer

Steven A. Ahrendt, John T. Chow, Li-Hua Xu, Stephen C. Yang, Claus F. Eisenberger, Manel Esteller, James G. Herman, Li Wu, P. Anthony Decker, Jin Jen, David Sidransky

Background: **Conventional cytologic analysis of sputum is an insensitive test for the diagnosis of non-small-cell lung cancer (NSCLC). We have recently demonstrated that polymerase chain reaction (PCR)-based molecular methods are more sensitive than cytologic analysis in diagnosing bladder cancer. In this study, we examined whether molecular assays could identify cancer cells in bronchoalveolar lavage (BAL) fluid.** ***Methods:*** **Tumor-specific oncogene mutations, CpG-island methylation status, and microsatellite alterations in the DNA of cells in BAL fluid from 50 consecutive patients with resectable (stages I through IIIa) NSCLC were assessed by use of four PCR-based techniques.** ***Results:*** **Of 50 tumors, 28 contained a p53 mutation, and the identical mutation was detected with a plaque hybridization assay in the BAL fluid of 39% (11 of 28) of the corresponding patients. Eight of 19 adenocarcinomas contained a K-ras mutation, and the identical mutation was detected with a mutation ligation assay in the BAL fluid of 50% (four of eight) of the corresponding patients. The p16 gene was methylated in 19 of 50 tumors, and methylated p16 alleles were detected in the BAL fluid of 63% (12 of 19) of the corresponding patients. Microsatellite instability in at least one marker was detected with a panel of 15 markers frequently altered in NSCLC in 23 of 50 tumors; the identical alteration was detected in the BAL fluid of 14% (three of 22) of the corresponding patients. When all four techniques were used, mutations or microsatellite instability was detected in the paired BAL fluid of 23 (53%) of the 43 patients with tumors carrying a genetic alteration.** ***Conclusion:*** **Although still limited by sensitivity, molecular diagnostic strategies can detect the presence of neoplastic cells in the proximal airway of patients with surgically resectable NSCLC. [J Natl Cancer Inst 1999;91:332–9].**

Lung cancer is the leading cause of cancer-related deaths among both males and females in the United States *(1)*. It was estimated that more than 170 000 new cases of primary lung cancer would be diagnosed in this country in 1998 and that more than 160 000 people will die of the disease *(1)*. Surgical resection remains the most effective form of treatment for non-small-cell lung cancer (NSCLC); however, at the time of diagnosis, more than 65% of all patients will have advanced disease that is no longer amenable to curative therapy. In addition, a large percentage of patients undergoing surgical resection ultimately die of recurrent NSCLC, demonstrating the frequent presence of occult metastatic disease at the time of diagnosis *(2)*. Attempts to improve lung cancer survival have focused on eliminating the cause (cigarette smoking), preventing the disease in high-risk groups (chemoprevention trials), diagnosing the disease at an early curable stage, and developing new adjuvant and neoadjuvant protocols. Each of these strategies has met with limited success.

Conventional cytologic analysis of sputum has not improved overall survival when added to a screening program of annual chest radiography for the early diagnosis of lung cancer *(3,4)*. Molecular techniques have identified tumor-specific oncogene mutations in cytologically negative sputum samples obtained from patients before the diagnosis of lung cancer *(5)*. However, the impressive sensitivities achieved by molecular techniques in this series were obtained in patients with dysplasia found by cytologic analysis of sputum.

In the present study, we examined the frequency of tumor-specific oncogene mutations, CpG-island methylation status, and microsatellite alterations in 50 consecutive, prospectively collected bronchoalveolar lavage (BAL) samples from patients with resectable NSCLC.

Materials and Methods

Sample Collection

Primary tumor, blood, and BAL fluid were collected prospectively from 50 consecutive patients undergoing surgical resection of NSCLC by a single surgeon at The Johns Hopkins Hospital or the Johns Hopkins Bayview Medical Center. BAL fluid was collected during flexible bronchoscopy performed at the time of pulmonary resection. Lavage was performed after guiding the bronchoscope into the segmental bronchus of the pulmonary lobe most likely to contain the tumor on the basis of earlier bronchoscopic and/or radiographic evaluation. Aliquots of 20 mL of warm saline were injected into the lobe until a volume of at least 30 mL was collected in a specimen trap. The BAL fluid was then transported to the laboratory on ice and centrifuged at 1800*g* for 10 minutes at 4 °C. The cell pellet was then collected and stored at −80 °C. Lymphocytes were collected from blood and used as a source of normal DNA. Tumor samples were promptly frozen at −80 °C after initial gross pathologic examination.

This research protocol was approved by the Joint Committee on Clinical Investigation of The Johns Hopkins School of Medicine in accord with an

Affiliations of authors: S. A. Ahrendt, P. A. Decker, Department of Surgery, Medical College of Wisconsin, Milwaukee; J. T. Chow, L.-H. Xu, C. F. Eisenberger, L. Wu, J. Jen, D. Sidransky (Department of Otolaryngology-Head and Neck Surgery), S. C. Yang (Department of Surgery), M. Esteller, J. G. Herman (The Oncology Center), The Johns Hopkins School of Medicine, Baltimore, MD.

Correspondence to: David Sidransky, M.D., Department of Otolaryngology-Head and Neck Surgery, The Johns Hopkins School of Medicine, 720 Rutland Ave., Baltimore, MD 21205-2196 (e-mail: dsidrans@welchlink.welch.jhu.edu).

See "Notes" following "References."

assurance filed with the U.S. Department of Health and Human Services. Written informed consent was obtained from all patients.

Pathologic stage was determined by the revised International System for Staging Lung Cancer *(6)*. Tumors were classified as peripheral, parenchymal, or central on the basis of preoperative radiographic studies, bronchoscopic or operative findings, and pathologic analysis. Peripheral tumors were located at or within 1 cm of the visceral pleura. Central tumors were visible on bronchoscopy or involved the main lobar bronchus upon pathologic examination. The remaining tumors were classified as parenchymal.

Portions of the primary tumor were cut into 7-μm sections, stained with hematoxylin–eosin, and examined by light microscopy. Additional 12-μm sections were cut and placed in a mixture of 1% sodium dodecyl sulfate and proteinase K (0.5 mg/mL) at 48 °C overnight. Tumors with a low neoplastic cellularity (<70%) were further microdissected to remove contaminating normal cells. The BAL cell pellet was also digested in 1% sodium dodecyl sulfate/proteinase K (0.5 mg/mL) as described above. DNA was then extracted from either type of sample with phenol/chloroform and precipitated with ethanol.

p53 Sequencing

A 1.8-kilobase fragment of the p53 gene (also known as TP53) (exons 5 through 9) was amplified from primary tumor DNA in all 50 patients by polymerase chain reaction (PCR) as described previously *(7,8)*. The PCR products were purified and sequenced directly by cycle sequencing (Amplicycle sequencing kit; The Perkin-Elmer Corp., Branchburg, NJ) by use of appropriate sequencing primers *(7,8)*. The products of the sequencing reactions were then separated by electrophoresis in 8 *M* urea/6% polyacrylamide gels, fixed, and exposed to film.

In addition, all 50 tumors were sequenced by use of the GeneChip® p53 assay (Affymetrix Inc., Santa Clara, CA) by the manufacturer's protocol. Exons 2 through 11 of the p53 gene from each tumor and the normal reference DNA were amplified as 10 amplicons in a single PCR. Each PCR mixture contained 250 ng of genomic DNA, 5 μL of the p53 primer set (Affymetrix Inc.), 10 U of AmpliTaq Gold (The Perkin-Elmer Corp.), PCR buffer II (The Perkin-Elmer Corp.), 2.5 m*M* $MgCl_2$, and all four deoxynucleoside triphosphates (dNTPs) (each at 0.2 m*M*) in a final volume of 100 μL. The reaction tubes were heated to 95 °C for 10 minutes and then subjected to 35 cycles of 95 °C for 30 seconds, 60 °C for 30 seconds, and 72 °C for 45 seconds and then to a final extension of 10 minutes at 72 °C. Forty-five microliters of amplified tumor and reference DNA was then fragmented with 0.25 U of fragmentation reagent (Affymetrix Inc.) at 25 °C for 18 minutes in a solution of 2.5 U of calf intestine alkaline phosphatase, 0.4 m*M* EDTA, and 0.5 m*M* Tris acetate (pH 8.2), followed by heat inactivation at 95 °C for 10 minutes.

The fragmented amplicons were then 3′ end labeled with fluorescein-labeled dideoxyadenosine 5′-monophosphate. Fifty microliters of fragmented DNA was incubated at 37 °C for 45 minutes in 100 μL containing 25 U of terminal deoxynucleotidyltransferase (Boehringer Mannheim Biochemicals, Indianapolis, IN), terminal deoxynucleotidyltransferase buffer, and 10 μ*M* fluorescein-N6-ddATP (where ddATP is dideoxyadenosine 5′-triphosphate), followed by heat inactivation at 95 °C for 5 minutes. The fluorescein-labeled sample was then hybridized in 0.5 mL containing 6× SSPE (1× SSPE = 0.15 *M* NaCl/10 m*M* sodium phosphate [pH 7.4]/1 m*M* EDTA), 0.05% Triton X-100, 1 mg of acetylated bovine serum albumin, and 2 n*M* control oligonucleotide F1 (Affymetrix Inc.) to the p53 probe array for 30 minutes at 45 °C. The probe array was washed four times with wash buffer A and then scanned by laser (HP GeneArray Scanner; Hewlett Packard, Wilmington, DE). The emitted light intensity is proportional to bound tumor DNA at each location on the probe array and was used by the GeneChip software to compare the nucleotide sequence of the tumor DNA to the reference sequence. Intensity patterns differing from the reference sample were quantified, and sites containing mutant bases were then displayed by the GeneChip software with an assigned score (range, 0–32) corresponding to the difference in intensity. The magnitude of the score also varied with the number of oligonucleotide probe sets present at a specific site. All sites were covered by two probe sets (one sense and one antisense). In addition, 300 known missense mutations in the p53 gene were also covered by an additional 14 probe sets. Scores exceeding an empirically determined threshold (a score of 13) were designated as a mutation. Samples with mutations detected by the GeneChip p53 assay that were not present on the direct sequencing gel were further confirmed by repeating the direct sequencing of the involved exon.

Oligonucleotide Hybridization

BAL samples from patients found to have a p53 mutation in their primary tumor were analyzed for the presence of tumor-specific p53 mutations with oligonucleotide plaque hybridization. The exon containing the p53 mutation was amplified for 35 cycles (94 °C for 30 seconds, 58 °C for 1 minute, and 70 °C for 1 minute) from the primary tumor, lymphocytes, and BAL fluid with primers containing *Eco*RI sites (exon 5, 4S [5′-GTAGGAATTCACTTGTGCCCTGACTT-3′] and 5ASECO [5′-ATCGAATTCAGACCTAAGAGCAAT-3′]; exon 6, 6SECO [5′-ATCGAATTCCCCAGGCCTCTGATT-3′] and 6ASECO [5′-ATCGAATTCGAGACCCAGTTGCAA-3′]; exon 7, 7SECO [5′-ATCGAATTCTCCCCAAGGCGCACT-3′] and 7ASECO [5′-ATCGAATTCGGTAAGAGGTGGGCC-3′]; exon 8, 8SECO [5′-ATCGAATTCAAATGGACAGGTAGA-3′] and 8ASECO [5′-ATCGAATTCTTGTCCTGCTTGCTT-3′]; and exon 9, 9SECO [5′-ATCGAATTCTCTGTTGCTGCAGAT-3′] and 9ASECO [5′-ATCGAATTCTGAGGTCACTCACCTGGA-3′]). The PCR products were then cloned into a λ bacteriophage vector (Stratagene Cloning Systems, La Jolla, CA) and amplified further in *Escherichia coli* cells *(7)*. Between 500 and 1000 clones were transferred to nylon membranes (NEN Research Products, Boston, MA) and hybridized with ^{32}P-end-labeled oligonucleotide probes specific for the p53 mutation identified in each patient's primary tumor as previously described (theoretical sensitivity = 1 : 1000 to 1 : 10 000) *(5)*. Hybridizing plaques signified the presence of a mutant p53 gene. After stripping, all filters were hybridized with a wild-type p53 oligonucleotide to identify plaques containing the inserted p53 fragments. All assays included both positive (tumor DNA with the mutation) and negative (tumor DNA with a different mutation and water alone) controls.

K-ras Mutation Ligation Assay

A 270-base-pair fragment containing exon 1 of the K-ras gene was amplified from the tumor and BAL fluid DNA of 27 patients with nonsquamous NSCLC (adenocarcinoma [n = 19], bronchoalveolar carcinoma [n = 6], adenosquamous carcinoma [n = 1], and large-cell carcinoma [n = 1]) as described previously *(9)*. The K-ras gene was also amplified from the tumor and BAL fluid DNA of 17 of the 23 patients with squamous cell cancer even though these tumors have been demonstrated previously to only rarely harbor K-ras gene mutations *(10)*. This fragment was used as the template for four separate mutation ligation assays (theoretical sensitivity = 1 : 100 to 1 : 1000) to detect all possible mutations at K-ras codon positions 12a, 12b, 13a, and 13b, as previously described *(11)*.

K-ras-Enriched PCR

A sensitive mutant-enriched PCR technique was also used to screen the tumor DNA of 27 patients with nonsquamous NSCLC and all 50 BAL samples for K-ras mutations *(12)*. In the first round of amplification, 1 μg of tumor DNA was amplified in 25 μL containing 60 ng of primers, 2.5 U of *Taq* polymerase (Sigma Chemical Co., St. Louis, MO), all four dNTPs (each at 0.2 m*M*), 50 m*M* KCl, 10 m*M* Tris–HCl (pH 8.3), and 1.5 m*M* $MgCl_2$ for 20 cycles (94 °C for 1 minute, 56 °C for 1 minute, and 72 °C for 1 minute). Five microliters of the first amplification products was then digested for 3 hours at 60 °C with *Bst*NI. The digested PCR products were then subjected to a second round of amplification (20 cycles) and digestion with *Bst*NI as described above. The products of the second *Bst*NI digestion were then separated by electrophoresis on nondenaturing polyacrylamide gels and stained with ethidium bromide [theoretical sensitivity = 1 : 10^4 to 1 : 10^6 *(12,13)*].

p16 Methylation-Specific PCR

Methylation-specific PCR was used to determine the methylation status of the CpG island of p16 in all 50 tumors and in the matched BAL fluid from 49 patients as described *(14)*. One microgram of tumor DNA or BAL-fluid DNA was modified with sodium bisulfite and precipitated with ethanol. The modified DNA was then amplified by use of both methylated- and unmethylated-specific primers as described *(14)*. PCR products were loaded directly onto nondenaturing 6%–8% polyacrylamide gels, stained with ethidium bromide, and visualized under UV illumination [theoretical sensitivity = 1 : 1000 *(14)*].

Microsatellite Analysis

Fifteen markers for microsatellite analysis were selected on the basis of previous work (our unpublished results), demonstrating frequent microsatellite al-

terations (expansion or deletion of a repeat unit) in NSCLC *(15)*. Tumor DNA and normal lymphocyte DNA from all 50 patients were examined with the entire panel of 15 microsatellite markers. BAL samples from 24 patients were analyzed with the 15-marker panel, whereas the remaining 26 BAL samples were examined only when the corresponding tumor demonstrated microsatellite instability. Oligonucleotides were obtained from Research Genetics (Huntsville, AL) or synthesized from sequences in the Genome Database (D3S1340, D3S1351, D8S321, D9S242, D11S488, D20S82, D20S85, CSFIR-1, and ACTβ-2) with the exception of the following loci: UT5307 (5′-GGATATAGCTGGCAATGGC-3′ [sense] and 5′-TCGGAATGCCTACTTCCCAG-3′ [antisense]), UT5320 (5′-ACCGACAGACTCTTGCCTC-3′ [sense] and 5′-TTGAGATGACCCTGAGACTG-3′ [antisense]), L17686 (5′-GCACCAATGCTCCAGAAATG-3′ [sense] and 5′-TCATGGTGCCATGATAGGAG-3′ [antisense]), L17835 (5′-TTGCACCACTATACTCCAGC-3′ [sense] and 5′-TCAGTTTAAGGTTCTCACCTG-3′ [antisense]), G29028 (5′-GCAGTGAGCTGAGATAATGC-3′ [sense] and 5′-TCACTAGCAGATGCGATAATG-3′ [antisense]), and G08460 (5′-TGGCGCTGATGCTCCACATTC-3′ [sense] and 5′-CTGGCTGACAGATAAAGCACT-3′ [antisense]). One marker from each primer pair was labeled with T4 polynucleotide kinase (New England Biolabs, Inc., Beverly, MA). PCR amplification was performed with 60 ng of DNA isolated from the tumor, normal lymphocytes, and BAL fluid as described above. Products were separated in 8% denaturing urea/polyacrylamide/formamide gels, followed by autoradiography [theoretical sensitivity = 1 : 100 to 1 : 200 *(16)*].

Statistical Analysis

Groups were compared by use of the Fisher exact test (two-tailed). All *P* values are two-sided.

Results

Primary Tumor

The clinical characteristics of 50 patients undergoing pulmonary resection for NSCLC are shown in Table 1. Twenty-eight patients had stage I disease (18 with stage IA and 10 with stage IB), 15 patients had stage II NSCLC (two with stage IIa and 13 with stage IIb), and seven patients had stage IIIa disease. Mean tumor size was 3.7 cm. The histologic type of the 50 tumors included squamous cell cancer (n = 23), adenocarcinoma (n = 19), bronchoalveolar carcinoma (n = 6), adenosquamous carcinoma (n = 1), and large-cell carcinoma (n = 1). p53 mutations were detected in tumors from 28 (56%) of 50 patients with NSCLC and were found statistically significantly (P = .014) more often in squamous cell cancer (19 [83%] of 23 patients) than in adenocarcinoma (eight [42%] of 19 patients). Nineteen (68%) of 28 p53 mutations in tumors were detected with direct cycle sequencing, and 22 (81%) of 27 mutations in tumors were detected with the p53 GeneChip. Mutations in the K-ras gene were identified in 10 (37%) of 27 patients with non-squamous-cell NSCLC, including eight (42%) of 19 patients with adenocarcinoma and two (33%) of six patients with bronchoalveolar carcinoma. Identical results were obtained with the mutation ligation assay and the K-ras mutation-enriched PCR. K-ras mutations were detected in two (12%) of 17 squamous cell cancers analyzed with the mutation ligation assay.

Methylated p16 alleles were present in 19 (38%) of 50 tumors tested. Microsatellite instability was identified in 23 (46%) of 50 tumors by use of the panel of microsatellite markers. These microsatellite "shifts" were seen in seven tumors (14%) with L17686; five tumors (10%) with D20S82; four tumors (8%) with UT5320; three tumors (6%) with L17835; two tumors (4%) with D8S321, D20S85, D9S242, and D3S1351; and one tumor (2%) with each of the following sequences, UT5307, G29028, D11S488, ACTβ-2, or G08460. The relationship between pathologic stage and cell type and the presence of p53 mutations, K-ras mutations, microsatellite instability, and methylated p16 alleles is shown in Table 1.

BAL Fluid

p53 mutant cells were detected in the BAL fluid of 11 (39%) of the 28 patients with p53 mutations in their primary tumor (Table 2 and Fig. 1). Four of the 12 tumors with K-ras mutations had the identical mutation detected in the BAL fluid (Fig. 2). The mutation ligation assay detected mutations in four samples, whereas the mutant-enriched PCR technique detected a mutation in only one sample (also detected by the ligation assay). None of the BAL samples from the 32 patients with a wild-type K-ras tumor contained a K-ras mutation when the mutation ligation assay was used. Only one of these 32 BAL samples contained a

Table 1. Frequency of p53 and K-ras mutations, any microsatellite instability, and p16 gene methylation status in the primary tumor from patients with resectable non-small-cell lung cancer

Parameter	n*	No. of patients/total No. (%)			
		p53 mutations	K-ras mutations†	Microsatellite instability‡	p16 methylation status
Pathologic stage§					
IA	18	11/18 (61)	2/15 (13)	7/18 (39)	8/18 (44)
IB	10	5/10 (50)	4/8 (50)	4/10 (40)	4/10 (40)
IIA	2	1/2 (50)	1/2 (50)	0/2	2/2 (100)
IIB	13	7/13 (54)	4/12 (33)	7/13 (54)	3/13 (23)
III	7	4/7 (57)	1/7 (14)	5/7 (71)	2/7 (29)
Total samples	50	28/50 (56)	12/44 (27)	23/50 (46)	19/50 (38)
Tumor cell type					
Squamous cell	23	19/23 (83)	2/17 (12)	14/23 (61)	9/23 (39)
Adenocarcinoma	19	8/19 (42)‖	8/19 (42)	7/19 (37)	9/19 (47)
Bronchoalveolar	6	0/6¶	2/6 (33)	1/6 (17)	1/6 (17)
Other	2	1/2 (50)	0/2	1/2 (50)	0/2
Total samples	50	28/50 (56)	12/44 (27)	23/50 (46)	19/50 (38)

*n = total number of patients with tumor of given stage or cell type.

†K-ras mutation analysis performed on 17 of 23 squamous cell cancers.

‡Number of tumors containing at least one microsatellite alteration in panel of 15 markers tested.

§Staging using the revised International System for Staging Lung Cancer *(6)*.

‖Two-sided P = .014 versus squamous cell cancers.

¶Two-sided P = .0008 versus squamous cell cancers.

Table 2. Frequency of tumor-specific p53 and K-ras mutations, any microsatellite instability, and p16 gene methylation status in the bronchoalveolar lavage (BAL) fluid from patients with resectable non-small-cell lung cancer

Parameter	n*	No. of BAL samples with detectable alteration/No. of BAL samples analyzed with corresponding alteration in primary tumor (%)				
		p53 plaque hybridization assay [1 : 1000]†	K-ras mutation ligation assay [1 : 200]†	Microsatellite instability [1 : 20]†	p16 methylation status [unknown]†	Any assay
Pathologic stage						
IA	18	1/11 (9)	1/2 (50)	1/7 (14)	5/8 (63)	7/17 (41)
IB	10	3/5 (60)	2/4 (50)	0/4	2/4 (50)	5/8 (63)
IIA	2	1/1 (100)	1/1 (100)	0/0	2/2 (100)	2/2 (100)
IIB	13	4/7 (57)	0/4	1/6 (17)	2/3 (67)	7/11 (64)
III	7	2/4 (50)	0/1	1/5 (20)	1/2 (50)	2/5 (40)
Total samples	50	11/28 (39)	4/12 (33)	3/22 (14)	12/19 (63)	23/43 (53)
Tumor cell type						
Squamous cell	23	6/19 (32)	0/2	3/13 (23)	5/9 (56)	11/21 (52)
Adenocarcinoma	19	4/8 (50)	4/8 (50)	0/7	7/9 (78)	11/17 (65)
Bronchoalveolar	6	0/0	0/2	0/1	0/1	0/4
Other	2	1/1 (100)	0/0	0/1	0/0	1/1 (100)
Total samples	50	11/28 (39)	4/12 (33)	3/22 (14)	12/19 (63)	23/43 (53)
Tumor site						
Peripheral	27	5/16 (31)	3/9 (33)	2/12 (17)	6/10 (60)	11/25 (44)
Parenchymal	13	2/8 (25)	1/2 (50)	0/6	3/5 (60)	6/11 (55)
Peripheral/parenchymal	40	7/24 (29)‡	4/11 (36)	2/18 (11)	9/15 (60)	17/36 (47)
Central	10	4/4 (100)	0/1	1/4 (25)	3/4 (75)	6/7 (86)
Total samples	50	11/28 (39)	4/12 (33)	3/22 (14)	12/19 (63)	23/43 (53)

*n = total number of patients with tumor of given stage, cell type, or site.

†Working limit of sensitivity defined as "proven" upper limit of sensitivity (i.e., detection of the molecular alteration in BAL fluid compared with presence of mutant alleles by quantitative p53 plaque hybridization assay) (*see* Table 3).

‡Two-sided $P = .032$ versus central tumors.

K-ras mutation when the mutant-enriched PCR was used. Methylated p16 alleles were detected in 12 of the 19 samples from patients with a methylated primary tumor (Fig. 3). None of the 30 BAL samples analyzed from patients with p16 unmethylated tumors contained methylated p16 alleles.

Tumor-specific microsatellite alterations (instability) were detected in only three (14%) of the 22 BAL fluids from patients with alterations present in the primary tumor (Fig. 4). We were unable to amplify a PCR product from one BAL sample with microsatellite instability detected in the corresponding tumor with marker D8S321. The BAL fluid from 24 patients was analyzed with the entire panel of 15 microsatellite markers. Microsatellite instability was present in 10 of these 24 tumors. However, tumor-specific microsatellite alterations were detected in the BAL fluid from only one (10%) of these 10 patients. Four (17%) of the 24 BAL samples contained a microsatellite alteration not present in the corresponding tumor.

One or more of the molecular assays were used to examine the BAL fluid from the 43 patients whose tumor displayed at least one molecular marker. Tumor-specific mutations or microsatellite instability was detected with at least one assay in 23 (53%) of these 43 BAL fluid samples.

Tumor location strongly influenced the ability to detect molecular alterations in BAL fluid. Tumor-specific p53 mutations were detected significantly more often in BAL fluid from patients with centrally located tumors than in BAL fluid from patients with tumors located in the parenchyma or peripherally (100% versus 29%; $P = .032$). Tumor size also played a significant role in detecting molecular alterations in the BAL fluid. Only one of 11 stage IA tumors had a detectable p53 mutation in the BAL fluid, and this tumor was 3.0 cm in size. The p53 plaque assay detected molecular alterations in the BAL fluid from patients with stage II tumors significantly more often than in the BAL fluid from patients with stage IA tumors (62% versus 9%; $P = .045$). The detection rate for tumor-specific p53 mutations in BAL fluid was similar between samples obtained from patients with squamous cell carcinoma (32%) and samples obtained from patients with adenocarcinoma (50%).

The ratio of mutant clones to total clones in the BAL fluid examined with the oligonucleotide plaque hybridization ranged from one in seven clones to one in 400 clones (Table 3). The ratio of mutant to normal clones was lower than one in 100 clones in all three positive BAL fluid specimens from patients with adenocarcinoma of the lung. In contrast, in five of the six patients with squamous cell carcinoma and in the one patient with large-cell carcinoma, the ratio of mutant to total clones in the BAL fluid was greater than one in 100 clones. The K-ras mutation ligation assay detected mutant cells to one in 200 cells. Methylated p16 alleles were detected in BAL fluid from 12 of the 19 patients with a methylated primary tumor. The p53 plaque assay was also positive (sample 971, one in seven clones; sample 1174, one in 400 clones) in just two of these samples (Table 3), giving a wide range for the sensitivity limit of the methylation assay. The sensitivity of the microsatellite assay was considerably lower as expected, detecting tumor-specific alterations in the BAL fluid only when the number of tumor cells was greater than one in 20 (5%) (Table 3).

Discussion

There is no effective test available to screen high-risk groups for lung cancer despite being the leading cause of cancer deaths among both men and women in the United States *(1,3,4)*. Conventional cytologic analysis of sputum was no more effective

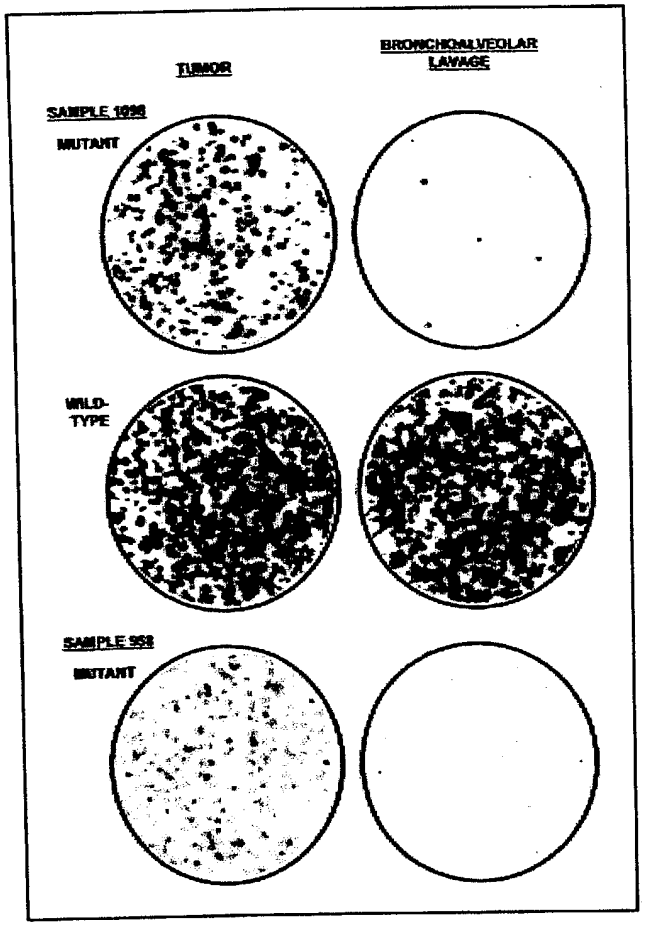


Fig. 1. Detection of mutant p53 alleles in bronchoalveolar lavage (BAL) fluid from patients with non-small-cell lung cancer. Nylon membranes were hybridized with both mutant-specific and wild-type oligonucleotides. A fragment of the p53 gene was amplified from DNA from the clinical samples as indicated, cloned, and transferred to nylon membranes. BAL fluid contained a few mutant p53 alleles (1 : 170) in sample 1098 but not in sample 958.

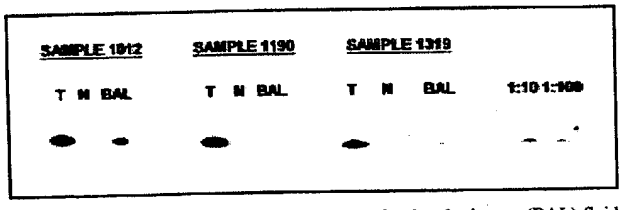


Fig. 2. Detection of mutant K-ras alleles in bronchoalveolar lavage (BAL) fluid with a mutation ligation assay. Mutant K-ras alleles (Cys^{12}) detected in BAL fluid in samples 1012 and 1319 but not in sample 1190. T = tumor; N = normal (wild-type). Lanes 1 : 10 and 1 : 100 contain a dilution of tumor DNA with K-ras mutation (Cys^{12}) with wild-type DNA from same patient (far right). Cys = cysteine (codon).

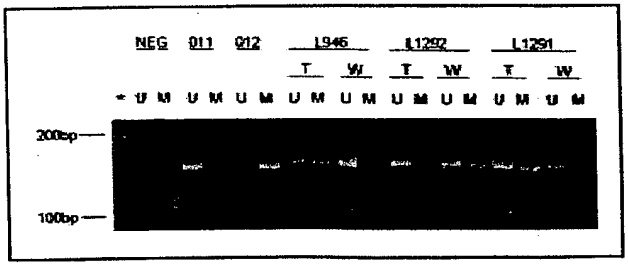


Fig. 3. Detection of methylated p16 alleles in tumor (T) and corresponding bronchoalveolar lavage (BAL) fluid (W) with methylation-specific polymerase chain reaction (PCR). Primer sets used for amplification are designated as unmethylated (U) or methylated (M). NEG shows the products of a negative control reaction devoid of DNA, and 011 and 012 show the products from positive control cell lines with unmethylated and methylated p16 alleles, respectively. Methylated p16 alleles were detected in the BAL fluid in sample L1292 but not in the BAL fluid in samples L946 and L1291. Additional bands represent nonspecific PCR products. * = marker lane; bp = base pairs.

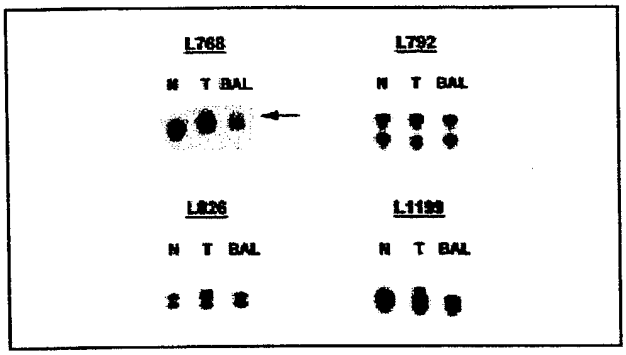


Fig. 4. Identification of microsatellite instability in clinical samples. Microsatellite instability (deletion of repeat units) in both tumor (T) and bronchoalveolar lavage (BAL) fluid in a patient with non-small-cell lung cancer (sample L768). Microsatellite alterations in primary tumor but not in BAL fluid (samples L792, L826, and L1199) were identified. N = normal lymphocyte DNA.

than annual chest radiographs in detecting lung cancer in several large prospective randomized trials *(3,4)*. Periodic screening with chest radiographs has not been demonstrated to decrease lung cancer mortality; however, their efficacy in lung cancer screening remains controversial *(17)*. In our study, four of the most promising molecular assays were used to evaluate BAL fluid obtained from 50 patients with resectable NSCLC and detected cancer cells in 53% of samples when the patient's primary tumor was positive for a molecular marker.

Cytologic analysis of sputum is currently used clinically to screen for or diagnose lung cancer; however, this test has several limitations. The majority of cancers detected by cytologic analysis are squamous cell cancers, whereas the most common cell type in NSCLC is adenocarcinoma *(18)*. Moreover, cytologic analysis is much more accurate at detecting central lesions than peripheral cancers, which also occur more commonly *(19)*. Bechtel et al. *(19)* reported on 51 patients with NSCLC detected by sputum cytologic testing. Eighty-six percent of the lesions were squamous cell cancer, and all but one of the tumors were centrally located and visible by bronchoscopic examination *(19)*. BAL has been used to improve the detection rate for peripheral lesions. Thirty-three percent of patients with peripheral NSCLC were detected by BAL in a recent series (tumor stage unknown in this group), and the sensitivity of BAL for diagnosing disseminated NSCLC may be as high as 88% *(20,21)*.

Several recent innovations, including the automated image cytometer or staining for heterogeneous ribonuclear protein overexpression, may improve the detection rate of cytologic techniques *(22,23)*. We elected to evaluate the ability of current molecular techniques at detecting tumor-specific molecular al-

Table 3. Comparison of p53 mutant alleles present in bronchoalveolar lavage (BAL) fluid and corresponding results of molecular assays with variable sensitivities

Sample No.	p53 mutation*	Cell type	Tumor site	No. of mutant alleles/ total No. of clones in BAL	Microsatellite instability†	K-ras mutation ligation assay†	p16 methylation assay†
971	Intron 6, A to G at 846	Squamous	Central	1/7	+		+
768	CGC to CAC at 175	Squamous	Peripheral	1/15	+		
847	ATC to TTC at 195	Squamous	Peripheral	1/20			
1011	GAG to TAG at 339	Large cell	Peripheral	1/24	−		
775	AGG to AGT at 249	Squamous	Parenchymal	1/45	−		
1216	CAG to TAG at 165	Squamous	Central	1/45	−		
1049	ACC to ATC at 253	Adenocarcinoma	Central	1/115			
1098	CGC to CTC at 158	Adenocarcinoma	Peripheral	1/170		+	
826	TGC to GGC at 176	Squamous	Central	1/200	−		
1174	GCC to CCC at 159	Adenocarcinoma	Parenchymal	1/400	−	−	−

*For all samples except 971, the numbers at right refer to nucleotide numbers of the published p53 complementary DNA sequence; for sample 971, the number at right refers to the intron 6 nucleotide number downstream from the 3′-end of exon 6 in the p53 gene.

†+ = positive; − = negative. If blank, the depicted molecular alteration was absent in the primary tumor and could not be evaluated.

terations in the BAL fluid from a group of patients with predominantly peripheral resectable NSCLC. In most cases, the entire BAL sample was used for molecular analysis to enable the completion of all four assays. One limitation to the current study is the unknown diagnostic sensitivity of cytologic analysis in this group of 50 patients to provide a direct comparison between cytologic and molecular techniques in detecting tumor cells in BAL fluid.

Oncogene or tumor suppressor gene mutations have been detected by several investigators *(5,12,13)* in the sputum of patients with primary adenocarcinoma of the lung. We identified K-ras and p53 mutations in eight of 10 sputum samples from patients who were participating in the Johns Hopkins Lung Project and who later developed adenocarcinoma *(5)*. However, each of these samples was selected from the 3% of samples collected in that trial that already had at least mild dysplasia upon cytologic examination. Several additional studies *(12,13)* have reported that K-ras mutations are detectable in the sputum of a high percentage of patients with adenocarcinoma when the primary tumor contains a K-ras mutation. Although K-ras mutations were frequently detected despite negative cytologic findings in most cases, each of these studies included a large number of patients with advanced unresectable lung cancer.

Several techniques have been used to detect K-ras mutations in sputum and stool samples; these techniques include plaque hybridization, the mutation ligation assay, and various mutant-enriched PCR techniques *(5,12,13,24)*. The mutation ligation assay is rapid, requires a single PCR amplification, and has a sensitivity approaching that of the plaque hybridization *(11)*. In this study, the mutation ligation assay was able to detect tumor cells with a sensitivity of at least one in 200 normal cells. Mutant-enriched PCR techniques have a reported sensitivity of at least one in 10^4; however, false-positive results are possible as the error rate of *Taq* polymerase or other thermostable enzymes is surpassed. Yakubovskaya et al. *(12)* used this technique to identify K-ras mutations in normal lung tissue from patients with lung cancer as often as from the primary cancer itself. In addition, 49% of patients with squamous cell cancer and 12% of control patients without cancer had K-ras mutations in their sputum. With a mutant-enriched PCR protocol designed to limit false-positive results, we were unable to improve upon the detection rate compared with the mutation ligation assay.

Although we detected K-ras mutant cells in 50% of patients with adenocarcinoma, molecular screening may still have a role in the early detection of patients with adenocarcinoma. Cytologic screening is notoriously poor in patients with adenocarcinoma, because adenocarcinoma cells are shed into the airway in low numbers *(18,19)*. In the Johns Hopkins Lung Project trial, the lung cancers detected by cytologic screening alone were almost exclusively of the squamous cell type *(18)*. In our study, all of the adenocarcinomas detected by the plaque assay contained fewer than 1% tumor cells in the BAL.

Tumor-specific microsatellite alterations have been used to identify tumor cells in the urine of patients with bladder cancer and in the serum of patients with head and neck and small-cell lung cancer *(25,26)*. In addition, tumor-specific microsatellite alterations have been demonstrated previously in the sputum of several patients, each with small-cell lung cancer and NSCLC *(15,16)*. We reported that this approach could detect approximately one cancer cell against a background of 200 normal cells, suggesting that the use of microsatellite alterations might prove useful as a diagnostic tool in NSCLC *(16)*. The reported frequency of microsatellite instability in NSCLC varies considerably from 2% to 55% in large part depending on the specific microsatellite loci examined *(27–29)*. We previously examined 73 microsatellite markers in 47 patients with NSCLC to find loci frequently altered in NSCLC and identified a panel of 12 markers that was able to detect at least one microsatellite alteration in 55% of the primary tumors (unpublished results). When we used the same panel and several additional markers found by other investigators to demonstrate microsatellite instability in NSCLC, we identified microsatellite alterations in a similar number of tumors (46%).

Microsatellite DNA markers are significantly more accurate than conventional cytology in identifying cancer cells in the [illegible] patients *(25)*. In contrast to molecular [illegible]e able to detect tumor-specific micro- in a small percentage (14%) of BAL ddition, microsatellite alterations not ing tumor were detected in 17% of the with the entire panel of 15 markers. A reported by Miozzo et al. *(15)*, who normal bronchial biopsy specimens from a primary lung cancer for micro- dentified microsatellite alterations that rimary tumor in 24% of these patients.

The results of the quantitative p53 plaque hybridization assay suggest that the threshold for detecting microsatellite alterations in BAL fluid was between a neoplastic cellularity of 5% and 7%. The percentage of tumor cells in the BAL fluid in this study was 10-fold to 100-fold lower than that in the urine of patients with bladder cancer and explains the low sensitivity of the microsatellite panel at detecting alterations in the BAL fluid.

Methylation-specific PCR has a sensitivity approaching one in 1000 in dilution experiments and may be useful in detecting clonal cell populations with methylated p16 genes in clinical samples *(14)*. In the current study, we were able to detect methylated p16 alleles in the BAL fluid from 63% of the patients with a methylated primary tumor. The quantitative p53 plaque assay was also positive in just two BAL samples with methylated p16 alleles, suggesting that the threshold for detecting methylated p16 alleles was between a tumor-to-normal cell ratio of one in seven and one in 400. Our results suggest that this approach may be useful in detecting cancers where the frequency of p16 inactivation by methylation is high *(30)*. However, unlike the primary tumor-specific oncogene mutations or microsatellite instability described above, p16 methylation is not necessarily specific. p16 methylation is a common molecular alteration and may not be limited to one neoplastic clone or may be present in other preneoplastic epithelial patches. Thus, although p16 methylation testing lends itself to prospective analysis without prior knowledge of the molecular alteration, it must be surveyed and carefully analyzed in normal control subjects and in cigarette smokers without cancer. Preliminary evidence from a pilot study on lung cancer suggests that p16 methylation may be detected in sputum in the absence of clinically evident cancer *(31)*. On the other hand, methylated p16 alleles were not observed in the BAL fluid from any of the patients with p16 unmethylated tumors in this study.

The molecular detection of lung cancer by testing BAL fluid (and certainly sputum) remains challenging. Although still limited by sensitivity, molecular approaches can detect tumor-specific mutations or molecular alterations in the proximal airway in resectable NSCLC. Further studies are needed to define the clinical significance of each of these mutations or alterations in patients at high risk of developing lung cancer. The more sensitive oligonucleotide plaque hybridization, mutation ligation, and p16 methylation assays detected these alterations in BAL fluid with similar frequency. Further improvements in the sensitivity of all of these assays may be achieved through enriching the epithelial cell component of the BAL fluid or sputum or by collecting multiple samples, an approach that has been successful at increasing the sensitivity of cytologic techniques. In addition, broad technical improvements in molecular detection assays and high through-put automation are necessary for eventual clinical implementation.

REFERENCES

(1) Landis SH, Murray T, Bolden A, Wingo PA. Cancer statistics, 1998 [published erratum appears in CA Cancer J Clin 1998;48:192]. CA Cancer J Clin 1998;48:6–29.

(2) Naruke T, Goya T, Tsuchiya R, Suemasu K. Prognosis and survival in resected lung carcinoma based on the new international staging system [published erratum appears in J Thorac Cardiovasc Surg 1989;97:350]. J Thorac Cardiovasc Surg 1988;96:440–7.

(3) Berlin NI, Buncher CR, Fontana RS, Frost JK, Melamed MR. The National Cancer Institute Cooperative Early Lung Cancer Detection Program. Results of the initial screen (prevalence). Early lung cancer detection: introduction. Am Rev Respir Dis 1984;130:545–9.

(4) Tockman MS. Survival and mortality from lung cancer in a screened population. Chest 1986;89:324S–325S.

(5) Mao L, Hruban RH, Boyle JO, Tockman M, Sidransky D. Detection of oncogene mutations in sputum precedes diagnosis of lung cancer. Cancer Res 1994;54:1634–7.

(6) Mountain CF. Revisions in the International System for Staging Lung Cancer. Chest 1997;111:1710–7.

(7) Brennan JA, Mao L, Hruban RH, Boyle JO, Eby YJ, Koch WM, et al. Molecular assessment of histopathological staging in squamous-cell carcinoma of the head and neck. N Engl J Med 1995;332:429–35.

(8) Boyle JO, Hakim J, Koch W, van der Riet P, Hruban RH, Roa A, et al. The incidence of p53 mutations increases with progression of head and neck cancer. Cancer Res 1993;53:4477–80.

(9) Ahrendt SA, Yang SC, Wu L, Westra WH, Jen J, Califano JC, et al. Comparison of oncogene mutation detection and telomerase activity for the molecular staging of non-small cell lung cancer. Clin Cancer Res 1997;3:1207–14.

(10) Rodenhuis S, Slebos RJ. Clinical significance of ras oncogene activation in human lung cancer. Cancer Res 1992;52 (9 Suppl):2665s–2669s.

(11) Jen J, Powell SM, Papadopoulos N, Smith KJ, Hamilton SR, Vogelstein B, et al. Molecular determinants of dysplasia in colorectal lesions. Cancer Res 1994;54:5523–6.

(12) Yakubovskaya MS, Spiegelman V, Luo FC, Malaev S, Salnev A, Zborovskaya I, et al. High frequency of K-ras mutations in normal appearing lung tissues and sputum of patients with lung cancer. Int J Cancer 1995;63:810–4.

(13) Mills NE, Fishman CL, Scholes J, Anderson SE, Rom WN, Jacobson DR. Detection of K-ras oncogene mutations in bronchoalveolar lavage fluid for lung cancer diagnosis [published erratum appears in J Natl Cancer Inst 1995;87:1643]. J Natl Cancer Inst 1995;87:1056–60.

(14) Herman JG, Graff JR, Myohanen S, Nelkin BD, Baylin SB. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. Proc Natl Acad Sci U S A 1996;93:9821–6.

(15) Miozzo M, Sozzi G, Musso K, Pilotti S, Incarbone M, Pastorino U, et al. Microsatellite alterations in bronchial and sputum specimens of lung cancer patients. Cancer Res 1996;56:2285–8.

(16) Mao L, Lee DJ, Tockman MS, Erozan YS, Askin F, Sidransky D. Microsatellite alterations as clonal markers for the detection of human cancer. Proc Natl Acad Sci U S A 1994;91:9871–5.

(17) Strauss GM, Gleason RE, Sugarbaker DJ. Screening for lung cancer. Another look; a different view. Chest 1997;111:754–68.

(18) Frost JK, Ball WC Jr, Levin ML, Tockman MS, Baker RR, Carter D, et al. Early lung cancer detection: results of the initial (prevalence) radiologic and cytologic screening in the Johns Hopkins study. Am Rev Respir Dis 1984;1:549–54.

(19) Bechtel JJ, Kelley WR, Petty TL, Patz DS, Saccomanno G. Outcome of 51 patients with roentgenographically occult lung cancer detected by sputum cytologic testing: a community hospital program [published erratum appears in Arch Intern Med 1994;154:1582]. Arch Intern Med 1994;154:975–80.

(20) de Gracia J, Bravo C, Miravitlles M, Tallada N, Orriols R, Bellmunt J, et al. Diagnostic value of bronchoalveolar lavage in peripheral lung cancer. Am Rev Respir Dis 1993;147:649–52.

(21) Poletti V, Romagna M, Allen KA, Gasponi A, Spiga L. Bronchoalveolar lavage in the diagnosis of disseminated lung tumors. Acta Cytol 1995;39:472–7.

(22) Payne PW, Sebo TJ, Doudkine A, Garner D, MacAulay C, Lam S, et al. Sputum screening by quantitative microscopy: a reexamination of a portion of the National Cancer Institute Cooperative Early Lung Cancer Study. Mayo Clin Proc 1997;72:697–704.

(23) Qiao YL, Tockman MS, Li L, Erozan YS, Yao SX, Barrett MJ, et al. A case–cohort study of an early biomarker of lung cancer in a screening cohort of Yunnan tin miners in China. Cancer Epidemiol Biomarkers Prev 1997;6:893–900.

(24) Sidransky D, Tokino T, Hamilton SR, Kinzler KW, Levin B, Frost P, et al. Identification of ras oncogene mutations in the stool of patients with curable colorectal tumors. Science 1992;256:102–5.

(25) Mao L, Schoenberg MP, Scicchitano M, Erozan YS, Merlo A, Schwab D,

et al. Molecular detection of primary bladder cancer by microsatellite analysis. Science 1996;271:659–62.

(26) Nawroz H, Koch W, Anker P, Stroun M, Sidransky D. Microsatellite alterations in serum DNA of head and neck cancer patients. Nat Med 1996;2:1035–7.

(27) Shridhar V, Siegfried J, Hunt J, del Mar Alonso M, Smith DI. Genetic instability of microsatellite sequences in many non-small cell lung carcinomas. Cancer Res 1994;54:2084–7.

(28) Fong KM, Zimmerman PV, Smith PJ. Microsatellite instability and other molecular abnormalities in non-small cell lung cancer. Cancer Res 1995; 55:28–30.

(29) Peltomaki P, Lothe RA, Aaltonen LA, Pylkkanen L, Nystrom-Lahti M, Seruca R, et al. Microsatellite instability is associated with tumors that characterize the hereditary non-polyposis colorectal carcinoma syndrome. Cancer Res 1993;53:5853–5.

(30) Liggett WH Jr, Sidransky D. Role of the p16 tumor suppressor gene in cancer. J Clin Oncol 1998;16:1197–206.

(31) Belinsky SA, Nikula KJ, Palmisano WA, Michels R, Saccomanno G, Gabrielson E, et al. Aberrant methylation of p16 ^{INK4a} is an early event in lung cancer and a potential biomarker for early diagnosis. Proc Natl Acad Sci U S A 1998;95:11891–6.

NOTES

Editor's note: Oncor, Inc., provided research funding for the study described in this article. Under an agreement between Oncor, Inc., and The Johns Hopkins University, Dr. Sidransky is entitled to a share of sales royalty received by The Johns Hopkins University from Oncor, Inc. Under that agreement, the university and Dr. Sidransky also have received Oncor stock, which, under university policy, cannot be traded until 2 years after the first commercial sales of the products related to this research. The terms of this arrangement have been reviewed and approved by the university in accordance with its conflict-of-interest policies.

Supported in part by Public Health Service grant K08CA76452-01 from the National Cancer Institute, National Institutes of Health, Department of Health and Human Services (S. A. Ahrendt); and by an American Association for Thoracic Surgery Research Scholarship (S. C. Yang).

Manuscript received June 22, 1998; revised November 24, 1998; accepted December 21, 1998.

EXHIBIT 3

Detection of K-*ras* Mutations of Bronchoalveolar Lavage Fluid Cells Aids the Diagnosis of Lung Cancer in Small Pulmonary Lesions[1]

Fumihiro Oshita,[2] Ikuo Nomura, Kouzo Yamada, Yuji Kato, Gaku Tanaka, and Kazumasa Noda

Division of Thoracic Disease, Kanagawa Cancer Center, Nakao 1–1–2, Asahi-ku, Yokohama 241-0815, Japan

ABSTRACT

An increased prevalence of K-*ras* oncogene mutation in lung adenocarcinoma has been shown by PCR-primer-introduced restriction with enrichment for mutation alleles (PCR-PIREMA) experiments. In the present study we investigated whether this method is useful for the diagnosis of lung cancer in small pulmonary lesions, which are difficult to diagnose cytologically as lung cancer by bronchoscopic examination. We examined bronchoalveolar lavage fluid (BALF) cells from 33 patients with single nodular pulmonary lesions of less than 2 cm in diameter (measured on chest computed tomography scans) for K-*ras* (codon 12) mutation, by PCR-PIREMA. Transbronchial fiberscopic examinations had not revealed lung cancer cytologically in any of the patients. The final diagnoses for the 33 lesions were 20 adenocarcinomas, 5 cases of focal fibrosis, 5 cases of pneumonia, 1 case of tuberculosis, 1 hamartoma, and 1 case of lymph node swelling. BALF cell lysates were amplified and digested with a restriction enzyme to detect the K-*ras* oncogene. Only the normal K-*ras* was observed after the first amplification and digestion for each of the 33 patients. Three amplifications and digestions were performed for each sample. We detected mutation of K-*ras* in BALF cells from 15 (75%) of 20 lung cancer patients and in cells from only 4 (31%) of 13 patients with nonmalignant lesions. The detection rate of the K-*ras* mutation in lung cancer was significantly greater than that in nonmalignant lesions ($P = 0.012$). Our results indicate that the detection of the codon 12 K-*ras* mutation in BALF cells by PCR-PIREMA aids the diagnosis of lung cancer in patients with small pulmonary lesions with negative cytological findings.

INTRODUCTION

Lung cancer is the leading cause of cancer deaths in Japan. To improve the prognosis of lung cancer patients, many oncologists have been trying to develop tests that will facilitate the earlier diagnosis and treatment of lung cancer and thereby decrease the mortality from this disease. Most early-stage lung cancers show no symptoms and are detected as an abnormal shadow on a chest roentgenogram or a chest CT[3] scan. Lung cancer appears as small nodules in the peripheral lung, and pathological or cytological diagnosis is essential for the diagnosis of lung cancer. Patients suspected of having lung cancer often undergo fiberscopic examination, with a tumor biopsy examination or a cytological approach. When a lesion is inaccessible to bronchoscopic biopsy or when the biopsy specimen is nondiagnostic, a diagnosis of cancer may be possible by cytological examination of the BALF, but this method is much less sensitive than the examination of a biopsy specimen (1). Cytological or pathological confirmation for small nodular lesions less than 2 cm in diameter is difficult; aggressive CT-guided aspiration cytology through the chest wall is often performed. However, many such lesions are resected without a diagnosis being made before the surgery.

Body fluids sometimes contain cells or cell debris bearing the oncogene mutations that characterize the related tumor, as has been shown for *ras* mutations in stool specimens from patients with colorectal tumors (2) and for *p53* mutations in urine from patients with bladder cancer (3). Similarly, mutations of K-*ras* that are associated with lung cancer have been detected in BALF cells (4). The clinical use of *ras* as a biomarker for lung cancer has been suggested by investigators who found *ras* mutations in stored sputum samples from patients later diagnosed with lung adenocarcinoma (5). In the largest study of *ras* mutations in human lung cancer, K-*ras* mutations in codon 12 predominated; they were found in 17% of 258 non-small cell lung cancer samples obtained by surgical resection, primarily (24%) in adenocarcinoma (6).

Cancer cells in BALF are always mixed with large numbers of genetically normal cells; therefore, the detection of *ras* mutations in BALF requires a sensitive assay such as PCR-PIREMA, which was developed to detect *ras* mutations (7, 8). It has been reported that the sensitivity and specificity of K-*ras* mutation detection in BALF samples for the diagnosis K-*ras* mutation-positive lung cancer were both 100% using PCR-PIREMA, and that this method detected K-*ras* mutations in BALF cells in 46% of adenocarcinomas of the lung (4). Therefore, we expected that we would frequently be able to detect the K-*ras* mutation in BALF cells from small lung lesions using the

Received 9/17/98; revised 11/24/98; accepted 12/8/98.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Supported in part by a Grant-in-Aid from the Ministry of Health and Welfare for the Second Term Comprehensive 10-year Strategy for Cancer Control.

[2] To whom requests for reprints should be addressed. Phone: 81-45-391-5761; Fax: 81-45-361-4692.

[3] The abbreviations used are: CT, computed tomography; BALF, bronchoalveolar lavage fluid; PCR-PIREMA, PCR-primer-introduced restriction with enrichment for mutation alleles.

Table 1 Patient characteristics

	No.
Total cases	33
Sex	
Male	18
Female	15
Age	
Mean	63
Range	45–79
Size of lesion[a] (cm)	
1.5–1.9	8
1.0–1.4	16
<1.0	9
Cytology[b]	
Negative	33

[a] The size of lesion was determined by chest CT.
[b] Cytological examination was performed by bronchoscopy.

Table 2 Final diagnosis of small pulmonary lesions

Diagnosis	No. of patients
Adenocarcinoma	20
Focal fibrosis	5
Pneumonia	5
Tuberculosis	1
Hamartoma	1
Swelling of lymph node	1

PCR-PIREMA method. We conducted a prospective study to determine whether the detection of K-*ras* mutation in BALF can aid the diagnosis of lung cancer in cases of a small pulmonary lesion that is cytologically negative on bronchoscopic examination.

MATERIALS AND METHODS

BALF Cell Collection. Between October 1995 and February 1998, patients with a nodular lesion of less than 2 cm in diameter in the peripheral lung that was subsequently diagnosed by biopsy specimen examination were enrolled in the present study. After chest roentgenography and CT, each patient underwent bronchoscopic examination to diagnose the cause of the lesion. Saline (50–100 ml) was injected into the target bronchus after transbronchial biopsy or brushing and lavage fluid specimens were obtained. One-half of the lavage fluid was used to make a cytological diagnosis, and the other half was used to test for K-*ras* mutation.

PCR-PIREMA Protocol. A modified PCR-PIREMA method was used to detect K-*ras* mutations in BALF cells (7). BALF cells (5×10^4) were washed and resuspended in 500 μl K-buffer [10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.5% Tween 20, and 100 μg/ml proteinase K]. The cell suspension was incubated at 55°C for 1 h and then heated at 94°C for 10 min to inactivate proteinase K. These cell lysates were stored at −20°C until used for PCR.

Briefly, PCR around K-*ras* codon 12 was performed by using a mismatched primer (F primer, 5′-ACTGAATATAAACTTGTGGTAGTTGGACCT-3′; R primer, 5′-ACTCATGAAAATGGTCAGAGAAACCTTTAT-3′) that introduced a *Bst*NI restriction site into the PCR products derived from normal alleles. *Bst*NI digestion of the PCR products left only the PCR products derived from mutant alleles intact, after which further PCR selectively amplified the mutant PCR products. The first PCR reaction mixtures contained 10 μl of cell lysate, 8 μM concentrations of each nucleotide, 0.8 mM $MgCl_2$, and 5′-mismatched primer to introduce a *Bst*NI restriction site flanking the K-*ras* exon 1. The first PCR products were digested with *Bst*NI. When mutated K-*ras* was not detected after digestion of the first PCR products, which had been amplified by 30 cycles of PCR, a fresh aliquot of the samples was amplified by 10 cycles of PCR, and, after *Bst*NI digestion, the samples was amplified twice more. The second PCR reaction mixture contained 10 μl of the digest of the first PCR products (diluted 1:100), 4 μM concentrations of each nucleotide, and 0.6 mM $MgCl_2$. The second PCR products also were digested with *Bst*NI. The first and second PCR reactions were performed for 10 and 20 cycles, respectively, at 94°C for 1 min, 55°C (for the first PCR) or 40°C (for the second PCR) for 2 min, and 74°C for 3 min. The digest of the second PCR products (diluted 1:100) was then amplified under standard PCR conditions (each nucleotide at 200 μM, 1.5 mM $MgCl_2$, 55°C annealing, 34 cycles) using the same primers followed by repeat *Bst*NI digestion; these products were then electrophoresed on 2.5% agarose gels and stained with ethidium bromide. A digestion-resistant 192-bp band indicated the presence of a K-*ras* codon 12 mutation. Each sample was subjected to the entire PCR-PIREMA process at least twice. Extensive measures were taken to prevent cross-contamination of samples. A normal control sample and a known mutation sample were included in all of the experiments.

Statistical Analysis. The χ^2 test was used to analyze the differences in the frequency of K-*ras* mutation between lung cancers and nonmalignant lesions.

RESULTS

Between October 1995 and February 1998, 56 patients with small nodular lesions in the peripheral lung that were less than 2 cm in diameter visited the Kanagawa Cancer Center. Fifteen of the patients were diagnosed with lung cancer cytologically by bronchoscopic examination, and eight of them did not receive a definite diagnosis and did not undergo surgical resection. The other 33 patients, for whom transbronchial examination did not reveal lung cancer cytologically although definite diagnosis was made later by surgical resection, entered the present study. The diameter of the lesion was less than 2 cm, and there was no lymph node swelling on chest CT for all of the 33 patients. The patient characteristics are shown in Table 1. The diameter of the lesion was between 1.5 cm and 1.9 cm for 8 patients, between 1.0 cm and 1.4 cm for 16 patients, and less than 1.0 cm for 9 patients. In all of the 33 cases, lung cancer was suspected based on the radiological findings, and the lesion was resected. The diagnoses were 20 cases of adenocarcinoma, 5 of focal fibrosis, 5 of pneumonia, 1 of tuberculosis, 1 of hamartoma, and 1 of lymph node swelling (Table 2). Data for K-*ras* mutation in BALF cells in four representative patients are shown in Fig. 1. Only normal K-*ras* was observed after the first amplification and digestion with the restriction enzyme for all of the four patients. We detected mutation in K-*ras* in the BALF cells from two patients after the third amplification and digestion with the

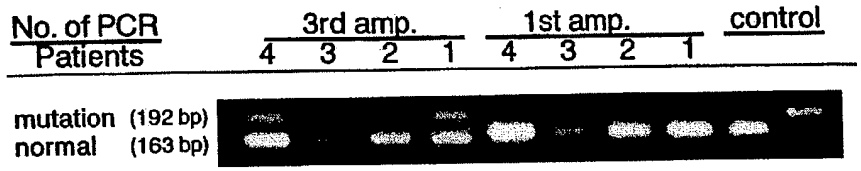


Fig. 1 PCR-PIREMA analysis of BALF cell lysates. *1st amp.*, PCR products of the first amplification (30 cycles) after digestion with *Bst*NI. *3rd amp.*, PCR products of the enriched screening step (three PCR amplifications followed by digestion with *Bst*NI: there were 10, 20, and 34 cycles in the first, second, and third amplifications, respectively.) The results for BALF cells from four representative patients are shown. Only the normal 163-bp band was observed after the first amplification in all of the cases, but a mutated K-*ras* 192-bp band was detected after the third amplification in cases 1 and 4. The diagnoses were adenocarcinoma (cases *1* and *4*) and pneumonia (cases *2* and *3*).

Table 3 Prevalence of K-*ras* codon 12 mutations in small pulmonary lesions by diagnosis

Detection of k-*ras* mutation in lung cancer was significantly higher than that in nonmalignancy ($P = 0.012$).

Diagnosis	No. of cases	
	Mutation (%)	Normal (%)
Lung cancer[a]	15 (75)	5 (25)
Nonmalignancy	4[b] (31)	9 (69)

[a] Adenocarcinoma (20).
[b] Focal fibrosis (3), pneumonia (1).

restriction enzyme. The lesions of the two patients with mutated K-*ras* were adenocarcinoma, and those of the other two patients, with normal K-*ras*, were pneumonia. The K-*ras* mutation results for all of the 33 patients are shown in Table 3. Only normal K-*ras* was observed after the first amplification and digestion for all of the 33 samples, and three sequential amplifications and digestions were performed for all of the samples. We detected mutation of K-*ras* in the BALF cells from 15 of the 20 lung cancer patients (75%) and in cells from only 4 of the 13 patients without a malignant lesion (31%). The nonmalignant lesions with a K-*ras* mutation were three cases of focal fibrosis and one of pneumonia. The detection rate of K-*ras* mutation in lung cancer was significantly higher than that in nonmalignant lesions ($P = 0.012$).

DISCUSSION

PCR-PIREMA is easily applied to BALF cells from patients undergoing diagnostic bronchoscopy as reported by Mills *et al.* (4, 7), and this method was reported to detect ras mutations at a higher rate than other methods. PCR-PIREMA has detected *ras* mutations in 46% of adenocarcinomas of the lung and detected one mutated allele in K-*ras* per 10^6 normal alleles. Most small lung cancers treated at our hospital are adenocarcinomas, and all of the lung cancer patients included in the present study had adenocarcinoma. Therefore, we examined whether detection of K-*ras* mutation using PCR-PIREMA aids the diagnosis of small pulmonary lesions.

Many genetic changes have been identified in lung cancer, but little is known about the chronology of their development. Some genetic changes may represent early activation events, whereas others are more likely to accompany late events related to invasion and metastasis. On the basis of the limited evidence, some investigators have suggested that in some human tumor types, including lung cancer, *ras* mutations may fall into the former category. The present study demonstrated that a codon 12 K-*ras* mutation was present in 15 (75%) of 20 lung cancer cases, which is higher than the previously reported frequencies. We analyzed very small lung cancer lesions, which in all 20 of the cases were adenocarcinoma. Our results may indicate that the K-*ras* mutation occurs more frequently in the early stage of adenocarcinoma compared with advanced adenocarcinoma. Analysis of the K-*ras* mutation in small resected lung cancers of less than 2 cm in diameter will verify this hypothesis.

Making cell lysates from BALF cells is easy and the PCR-PIREMA assay is rapid, nonradioactive, and readily adaptable to processing large numbers of clinical samples. The assay can also be used to detect all of the activating mutations of K-*ras*. However, three PCR amplifications are required to detect one mutant allele in 10^6 normal alleles, and there is a high misincorporation rate of Taq polymarase. One error occurs per 10^4 bases under standard PCR conditions. To minimize the PCR error, we used the protocol reported by Mills *et al.* (4), but we decreased the cycle number for the first and second amplifications.

Fifteen of 20 adenocarcinomas had a mutated K-*ras* in the BALF cells, which is significantly higher than the mutated K-*ras* frequency that we saw in patients with nonmalignant lesions (4 of 13). Therefore, the detection of mutated K-*ras* in BALF cells may indicate lung cancer of the target lesion in spite of a pathological or cytological negative finding. K-*ras* mutation was detected in BALF cells from four patients with nonmalignant lesions: three had focal fibrosis, and one had pneumonia. Atypia of cells was observed in a lesion removed from one the focal fibrosis patients, and active inflammation was observed in the lesion from the patient with pneumonia. Therefore, a K-*ras* mutation may present in nonmalignant lesions such as atypia and active inflammation, and the detection of a K-*ras* mutation does not always correlate with cytological detection of cancer cells. Nevertheless, the frequency of a K-*ras* mutation in BALF cells is about 2.5 times greater in cases of lung cancer, and, therefore, the K-*ras* mutation in BALF cells is a clinically useful biomarker for lung cancer. In conclusion, detection of a K-*ras* mutation in BALF cells using PCR-PIREMA aids the diagnosis of lung cancer in patients with small pulmonary lesions. Our results suggest that surgical resection should be considered for patients with a K-*ras* mutation in their BALF cells even if the cytological finding is negative.

REFERENCES

1. Minna, J. D. Cancer of the lung. *In:* V. Devita, Jr., S. Hellman, and S. A. Rosenberg (eds), Cancer: Principles and Practice of Oncology, pp. 591–705. Philadelphia: Lippincott, 1989.

2. Sidransky, D., Tokino, T., Hamilton, S. R., Kinzler, K. W., Levin, B., Frost, P., and Vogelstein, B. Identification of *ras* oncogene mutations in the stool of patients with curable colorectal tumors. Science (Washington DC), *256:* 102–105, 1992.

3. Sidransky, D., Von Eschenbach, A., Tsai, Y. C., Jones, P., Summerhayes, I., Marshall, F., Paul, M., Green, P., Hamilton, S. R., Frost, P., and Vogelstein, B. Identification of *p53* gene mutations in bladder cancers and urine samples. Science (Washington DC), *252:* 706–709, 1991.

4. Mills, N. E., Fishman, C. L., Scholes, J., Anderson, S. E., Rom, W. N., and Jacobson, D. R. Detection of K-*ras* oncogene mutations in bronchoalveolar lavage fluid for lung cancer diagnosis. J. Natl. Cancer Inst., *87:* 1056–1060, 1995.

5. Mao, L., Hruban, R. H., Boyle, J. O., Tockman, M., and Sidransky, D. Detection of oncogene mutations in sputum precedes diagnosis of lung cancer. Cancer Res., *54:* 1634–1637, 1994.

6. Rodenhuis, S., and Slebos, R. J. Clinical significance of *ras* oncogene activation in human lung cancer. Cancer Res., *52* (Suppl. 9): 2665s–2669s, 1992.

7. Mills, N. E., Fishman, C. L., Rom, W. N., Dubin, N., and Jacobson, D. R. Increased prevalence of K-*ras* oncogene mutations in lung adenocarcinoma. Cancer Res., *55:* 1444–1447, 1995.

8. Jacobson, D. R., and Mills, N. E. A highly sensitive assay for mutation *ras* genes and its application to the study of presentation and relapse genotypes in acute leukemia. Oncogene, *9:* 553–563, 1994.

EXHIBIT 4

1810–1818 Nucleic Acids Research, 1999, Vol. 27, No. 8

Nucleotide analogs facilitate base conversion with 3′ mismatch primers

Joseph P. Day, Donald Bergstrom[1], Robert P. Hammer[2] and Francis Barany*

Department of Microbiology, Box 62, Hearst Microbiology Research Center, Strang Cancer Prevention Center, Joan and Sanford I. Weill Medical College of Cornell University, 1300 York Avenue, New York, NY 10021, USA, [1]Department of Medicinal Chemistry and Molecular Pharmacology, Purdue University, West Lafayette, IN 46208, USA and [2]Department of Chemistry, Louisiana State University, Baton Rouge, LA 70803, USA

Received as resubmission January 19, 1999; Revised and Accepted February 18, 1999

ABSTRACT

We compared the efficiency of PCR amplification using primers containing either a nucleotide analog or a mismatch at the 3′ base. To determine the distribution of bases inserted opposite eight different analogs, 3′ analog primers were used to amplify four different templates. The products from the reactions with the highest amplification efficiency were sequenced. Analogs allowing efficient amplification followed by insertion of a new base at that position are herein termed 'convertides'. The three convertides with the highest amplification efficiency were used to convert sequences containing C, T, G and A bases into products containing the respective three remaining bases. Nine templates were used to generate conversion products, as well as non-conversion control products with no base change. We compared the ability of natural bases to convert specific sites with and without a preconversion step using nucleotide analog primers. Conversion products were identified by a ligation detection reaction using primers specific for the converted sequence. We found that conversions resulting in transitions were easier to accomplish than transversions and that sequence context influences conversion. Specifically, primer slippage appears to be an important mechanism for producing artifacts via polymerase extension of a 3′ base or analog transiently base paired to neighboring bases of the template. Nucleotide analogs could often reduce conversion artifacts and increase the yield of the expected product. While new analogs are needed to reliably achieve transversions, the current set have proven effective for creating transition conversions.

INTRODUCTION

Highly sensitive assays that detect low abundance mutations rely on PCR to amplify the target sequence. However, a non-selective PCR strategy will amplify both mutant and wild-type alleles with approximately equal efficiency, resulting in low abundance mutant alleles comprising only a small fraction of the final product. If the mutant sequence comprises <25% of the amplified product, it is unlikely that DNA sequencing will be able to detect the presence of such an allele. Although it is possible to accurately quantify low abundance mutations by first separating the PCR products by cloning and subsequently probing the clones with allele-specific oligonucleotides (ASOs) (1–3), this approach is time consuming. In contrast to the above, allele-specific PCR methods can rapidly and preferentially amplify mutant alleles. For example, multiple mismatch primers have been used to detect H-*ras* mutations at a sensitivity of one mutant in 10^5 wild-type alleles (4) and claims as high as one mutant in 10^6 wild-type alleles have been reported (5,6). However, careful evaluation suggests these successes are limited to allele-specific primers discriminating through 3′ purine·purine mismatches. For the more common transition mutations, the discriminating mismatch on the 3′ primer end (i.e. G:T or C:A mismatch) will be removed in a small fraction of products by polymerase error during extension from the opposite primer on wild-type DNA. Thereafter, these error products are efficiently amplified and generate false positive signal. One strategy to eliminate this polymerase error problem is to deplete wild-type DNA early in PCR.

Several investigators have explored selective removal of wild-type DNA by restriction endonuclease digestion in order to enrich for low abundance mutant sequences. These restriction fragment length polymorphism (RFLP) methods detect approximately one mutant in 10^6 wild-type or better by combining the sensitivity of polymerase with the specificity of restriction endonucleases. One approach has used digestion of genomic DNA followed by PCR amplification of the uncut fragments (RFLP–PCR) to detect very low level mutations within restriction sites in the H-*ras* and p53 genes (7,8). Similar results have been obtained by digestion following PCR and subsequent amplification of the uncleaved DNA now enriched for mutant alleles (PCR–RFLP) (9–11). Although sensitive and rapid, RFLP detection methods are limited by the requirement that the location of the mutations must coincide with restriction endonuclease recognition sequences. To circumvent this limitation, primers that introduce a new restriction site have been employed in 'primer-mediated RFLP' (12–17). However, subsequent investigators have demonstrated

*To whom correspondence should be addressed. Tel: +1 212 746 4288; Fax: +1 212 746 8588; Email: barany@mail.med.cornell.edu

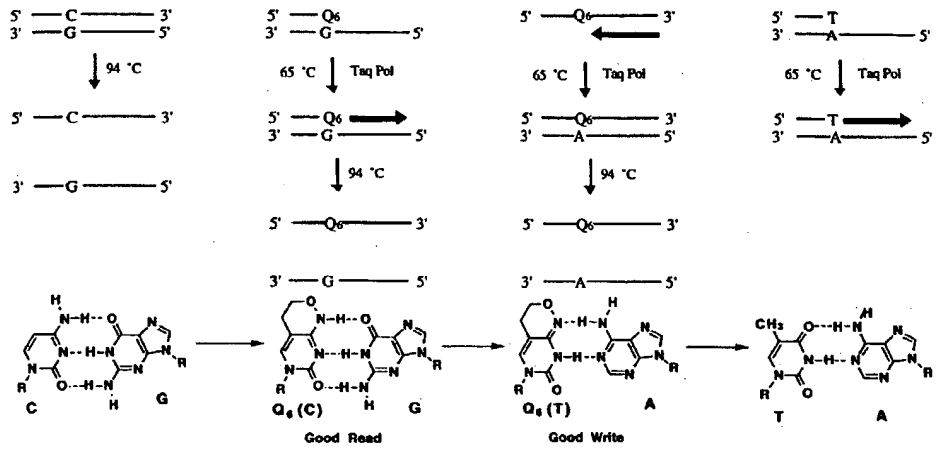


Figure 1. Conversion facilitated by nucleotide analog preconversion. A C:G base pair in a sequence is targeted for conversion to a T:A base pair. Rather than using a 3′ natural base mismatch primer to attempt direct conversion, a nucleotide analog (Q_6) primer is used for preconversion. The Q_6 analog reads the G base well and allows polymerase to efficiently extend from the 3′ Q_6 primer. During PCR, the reverse primer anneals to the Q_6 PCR product and is extended by polymerase to synthesize the opposite strand. When polymerase reaches the Q_6 analog in the template, polymerase writes A (or G; not shown) opposite the analog and continues synthesis of the strand. After a few cycles, a pool of products is made with degenerate sequence opposite the analogs. A natural base primer is then added to selectively amplify the products having the desired base change.

that errors are produced at the very next base by polymerase extension from primers having 3′ natural base mismatches (18–20). Such templates fail to cleave during restriction digestion and amplify as false positives that are indistinguishable from true positive products extended from mutant templates.

Use of nucleotide analogs may reduce errors resulting from polymerase extension and improve base conversion fidelity. Nucleotide analogs that are designed to base pair with more than one of the four natural bases herein are termed 'convertides'. Base incorporation opposite different convertides has been tested (21). For each analog, PCR products were generated using *Taq* polymerase and primers containing an internal nucleotide analog. The products generated showed a characteristic distribution of the four bases incorporated opposite the analogs. Of significance, these products retained the original sequence at all natural base positions. Convertides readily form degenerate amplification products by virtue of their ability to assume different hydrogen bonding patterns through either tautomeric shift (22), bond rotation (23) or base stacking (24,25). Thus, PCR primers containing convertides may be used to facilitate base conversion. In principle, using the 6*H*,8*H*-3,4-dihydropyrimido[4,5-*c*][1,2]-oxazine-7-one analog (Q_6), which is known to exhibit both the C-like and T-like tautomeric forms at the 3′-end of the primer (22), a C-G base pair may be converted to a T-A base pair (Fig. 1). Due to the better geometry, DNA polymerases may 'read' or extend better from a Q_6·G pair than a T·G mismatch (wobble base pair). Similarly, DNA polymerases may 'write' or incorporate both G and A bases opposite Q_6 (26), whereas A is always inserted opposite a T base. Thus, the Q_6 analog primer serves as an intermediary, providing a 'preconversion' step before a natural base primer is added to selectively amplify the desired product from the degenerate pool. While nucleotide analogs have great potential, they have not been tested in high sensitivity assays.

We synthesized several PCR primers containing one of eight different nucleotide analogs at the 3′-end (Fig. 2). PCR extension

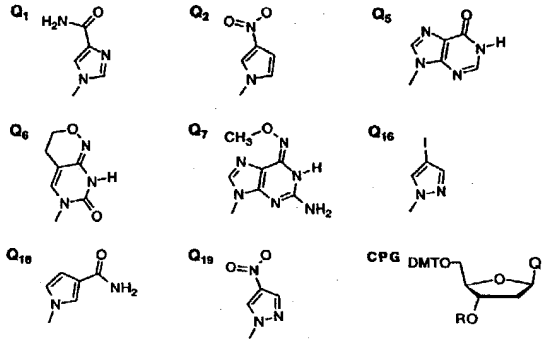


Figure 2. Nucleotide analogs used in PCR primers. In the final deprotected oligonucleotide, the name of the nucleoside containing the base analog shown is: Q_1, 1-(2′-deoxy-β-D-ribofuranosyl)imidazole-4-carboxamide; Q_2, 1-(2′-deoxy-β-D-ribofuranosyl)-3-nitropyrrole; Q_5, 2′-deoxyinosine; Q_6, 6-(2′-deoxy-β-D-ribofuranosyl)-6*H*,8*H*-3,4-dihydropyrimido[4,5-*c*][1,2]oxazine-7-one; Q_7, 2-amino-7-(2′-deoxy-β-D-ribofuranosyl)-6-methoxyaminopurine; Q_{16}, 1-(2′-deoxy-β-D-ribofuranosyl)-4-iodopyrazole; Q_{18}, 1-(2′-deoxy-β-D-ribofuranosyl)pyrrole-3-carboxamide; Q_{19}, 1-(2′-deoxy-β-D-ribofuranosyl)-4-nitropyrazole. Base analogs (Q) are attached to the 1′ position of deoxyribofuranose. The nucleoside analogs are attached to the controlled pore glass (CPG) column via a succinoyl linker (R, linker to CPG). The oligonucleotide is synthesized from the 5′-hydroxyl after removal of the dimethoxytrityl (DMT) protecting group, placing the analog at the 3′-end. After cleavage from the CPG column and deprotection, the oligonucleotide is extended by polymerase from the 3′ base analog hydroxyl group (R = H).

efficiency and fidelity were measured and the mutations in PCR products identified by sequencing and ligation detection reaction (LDR) (27–29). We found that primer-mediated RFLP–PCR using natural base 3′ mismatch primers is prone to high levels of misextension errors. Specific misextension errors in each reaction

were quantified in the range 0.1–100% using LDR (30). However, conversion fidelity could be significantly improved if preconversion with 3′ convertide primers was performed.

MATERIALS AND METHODS

Oligonucleotide synthesis

Oligonucleotides were synthesized at the 0.2 µmol scale by cyanoethyl phosphoramidite chemistry on an Applied Biosystems 394 DNA synthesizer. Standard 500 Å CPG columns and reagents (Applied Biosystems) were used with the following exceptions. Oligonucleotides 50 bases in length were synthesized using wide pore 1000 Å CPG columns (Applied Biosystems). Oligonucleotides with fluorescent dye FAM at the 5′-terminus were synthesized using FAM phosphoramidite (Applied Biosystems) with a 15 min coupling step. Oligonucleotides with 5′ phosphate were synthesized using phosphorylation reagent (Glen Research) with a 15 min coupling step. Oligonucleotides with 3′ blocking group were synthesized using 3′-Spacer CPG columns (Glen Research). Oligonucleotides with the 3′ nucleotide analogs 2′-deoxyinosine (Q_5), 6-(2′-deoxy-β-D-ribofuranosyl)-6*H*,8*H*-3,4-dihydropyrimido[4,5-c][1,2]oxazine-7-one (Q_6) and 2-amino-7-(2′-deoxy-β-D-ribofuranosyl)-6-methoxyaminopurine (Q_7) were synthesized using 2′-deoxyinosine-CPG, dP-CPG and dK-CPG, respectively (Glen Research) (Fig. 1). The oligonucleotide primers containing Q_1, Q_2 and Q_{18} at the 3′-position were synthesized from Q_1-, Q_2- and Q_{18}-derived CPG synthesized from Q_1 (31), Q_2 (24) and Q_{18} (25) by the method of Pon *et al.* (32). Details of the synthesis of the iodopyrazole (Q_{16}) and nitropyrazole (Q_{19}) nucleosides will be reported separately.

PCR polymerases and buffers

The DNA polymerases used were Ampli*Taq*, Ampli*Taq* Stoeffel Fragment, Ampli*Taq* Fluorescent Sequencing (Applied Biosystems), Vent and Vent(exo–) (New England Biolabs) and Expand polymerase mix (*Taq* and *Pfu* polymerase mixture, in Expand High Fidelity kit; Boehringer Mannheim). The commercially available PCR buffers used were supplied in the Ampli*Taq* and Expand High Fidelity kits. An alternative buffer, CiNF, is described elsewhere (33). Briefly, CiNF reactions contain 20 mM citrate, pH 7.6, 200 µg/ml bovine serum albumin, 2.5 mM $MgCl_2$, 200 µM dNTP (each) and either 16 mM $(NH_4)_2SO_4$ or 50 mM potassium acetate, 10% formamide, primers and template DNA. All PCR and LDR reactions described below were performed under paraffin oil.

Mismatch extension efficiency

Primers containing natural bases and nucleotide analogs were used in PCR to measure the efficiency of product formation from synthetic duplex p53 exon 7 templates having *Msp*I (CCGG), *Taq*I (TCGA), *Hha*I (GCGC) or *Tai*I (ACGT) sites at the *Msp*I position containing codon 248. The primers hybridized to wild-type sequence on either side of the *Msp*I site with the 3′-ends of the primers extending one base into the site on each side (Fig. 3A). Eight different analogs and the four natural bases were tested in parallel reactions on each of the four synthetic templates. PCR was performed using *Taq* Stoeffel Fragment or *Taq* Fluorescent Sequencing polymerases with the buffer supplied for each polymerase. We used 10 pmol of each primer and 20 fmol of duplex template, 0.2 mM each dNTP and 4 mM $MgCl_2$. Parallel reactions underwent 10, 20, 30, 40 and 50 PCR cycles of 94°C for 15 s, 65°C for 1 min. Efficiency and yield were determined from samples run on 3% agarose gels and stained with ethidium bromide.

Mismatch conversion product sequencing

Products most efficiently amplified by each analog were diluted 1000-fold in water. The diluted DNA products were reamplified for 20 cycles of 94°C for 15 s, 65°C for 2 min using the same polymerase and buffer as in the previous PCR, but with the addition of 10 pmol of 'zipcode'-containing primers p53zip248 and p53zip248R (Fig. 3A). Zipcode sequences are oligonucleotides with no known sequence similarity to DNA sequences in any organism. Amplification with zipcode primers is intended to specifically amplify the zipcode-containing products of the previous PCR, i.e. only converted DNA (containing zipcodes) and not the nearly identical unconverted DNA (lacking zipcodes) will be amplified. Conversion products were run on 3% agarose gels and bands of the expected size excised. DNA was extracted from the gel slices by centrifugation in a 235C microcentrifuge (Fisher) for 30 min through a 0.45 µm HVLP filter (Millipore). The conversion product was dried and resuspended in ABI Dye Terminator Cycle Sequencing reaction mix with one of the zipcode primers according to kit instructions (Applied Biosystems). An equal volume (3 µl each) of sequencing reaction was combined with dye mix consisting of 83% formamide (Eastman), 4 mM EDTA and 8 mg/ml Blue Dextran (Sigma). Samples were electrophoresed on a 7 M urea–10% acrylamide gel (19:1 bis, 0.6× TBE in gel and running buffer) in an ABI 373 DNA Sequencer. Data were analyzed using ABI 373A DNA Sequencer Data Analysis software v.1.2.0.

Conversion product identification

Conversion fidelity was tested using nine different synthetic templates, with and without preconversion using three primers containing Q_5, Q_6 and Q_7 (see Oligonucleotide synthesis). Preconversion PCR was performed with 3′ analog primers prior to adding the desired natural base primers, in an effort to avoid mismatch primer extension. The 50 bp duplex DNA templates contained the wild-type p53 sequence surrounding codon 248 (Fig. 3B), except for the bases corresponding to the *Msp*I site (CCGG). The following sequences were substituted at the *Msp*I position: 1) CCGG (wild-type); 2) CTGG; 3) CGGG; 4) CAGG; 5) TCGA; 6) GCGC; 7) ACGT; 8) ACGT; 9) GCGC. Preconversion was performed with hot start using 50 fmol/µl p53-248Q_N and p53-248Q_NR primers and Vent(exo–) in CiNF buffer and 10 fmol/µl of duplex template. Preconversion used two PCR cycles of 94°C for 15 s, 55°C for 1 min, 60°C for 1 min. Product was stored at 4°C. Conversion reactions were started with 1 µl of preconversion reaction containing the same polymerase and buffer, but no additional template. Each reaction required 10 pmol of each primer, using one of the four pairs p53zip248N and p53zip248NR (N = C, T, G or A). Parallel conversion reactions with no preconversion were initiated with a hot start by adding 10 fmol of synthetic duplex template instead of preconversion reaction aliquot. PCR cycles were as follows: five cycles of 94°C for 15 s, 55 + 1°C/cycle for 1 min, 60°C for 1 min; 20 cycles of 94°C for 15 s, 60°C for 2 min. A final extension was performed at 60°C for 5 min. Polymerase was inactivated by freezing and

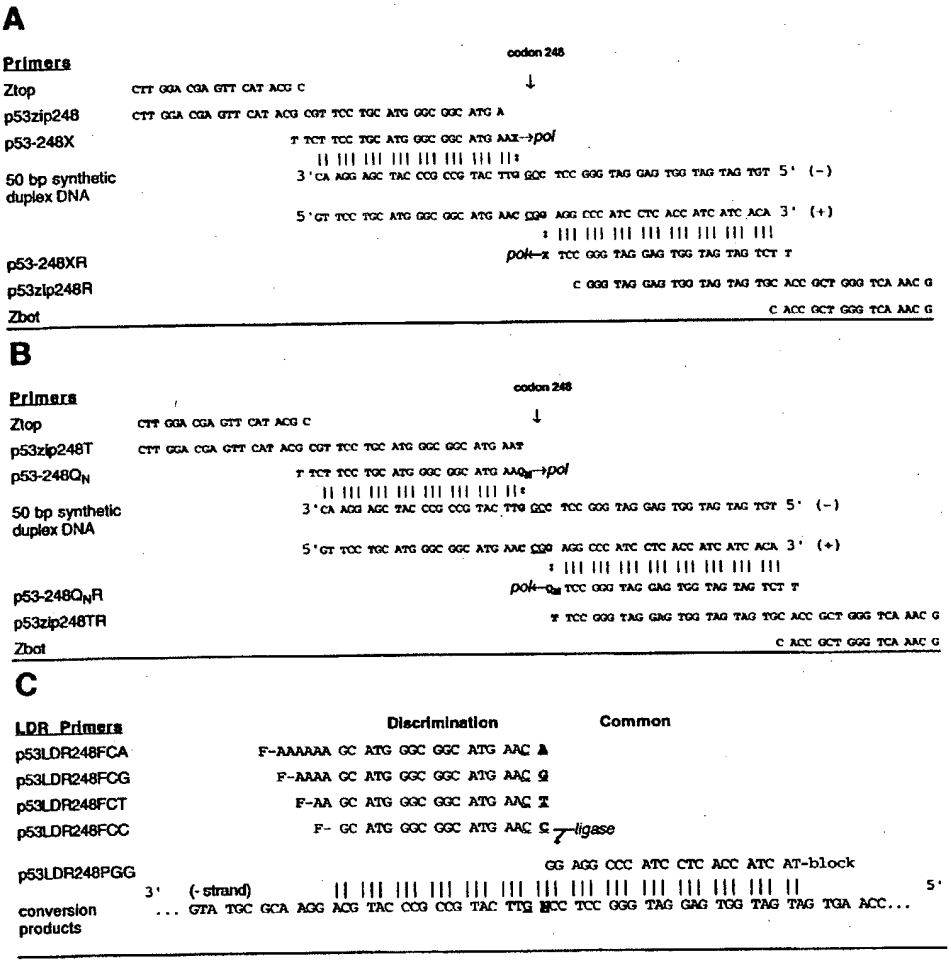


Figure 3. Primers used in mismatch extension and PCR/LDR. Complimentary (–strand) sequences are shown in reverse orientation (3′→5′), e.g. reverse strand primers (names ending in R). (**A**) One of nine different synthetic 50 bp duplex templates is shown melted with primers aligned to complementary sequence. Primer extension was performed using 3′ natural base and nucleotide analog primers (p53-248X and p53-248XR). Some extension products were reamplified using truncated zipcode primers p53zip248 and p53zip248R and sequenced using one of the zipcode primers (Ztop or Zbot). (**B**) Preconversion was performed on nine different 50 bp synthetic duplex templates using 3′ nucleotide analog primers, e.g. p53-248Q_6 and p53-248Q_6R. Conversion, with or without preconversion, was performed using primers containing the 3′ natural base, e.g. primers p53zip248T and p53zip248TR. These conversion products were reamplified using zipcode primers and identified by LDR. (**C**) LDR primer sets were designed to identify specific base changes in conversion products. LDR primers anneal in competition with each other to conversion products. Perfectly complementary upstream and downstream LDR primers with no overlap or gap ligate with high specificity. Discrimination primers had different length 5′ tails to allow specific product separation on an acrylamide gel. Shown are a set of primers used to identify PCR error products in non-conversion of wild-type template.

thawing twice. Products were diluted 10× in water and reamplified by adding 1–20 µl of Expand polymerase and buffer mix. PCR was performed for 20 cycles (30 cycles for low yield reactions) of 94°C for 15 s, 65°C for 2 min using 12 pmol of zipcode primers Ztop and Zbot (Fig. 3). LDR was performed as described below to identify the conversion products generated.

Ligase detection reaction

Ligase detection reactions were performed in standard LDR buffer (25 mM Tris pH 7.6, 12 mM $MgCl_2$, 65 µg/ml bovine serum albumin, 100 mM KCl and 10 mM DTT). Each 20 µl reaction contained ~500 fmol of dsDNA (1 µl of PCR sample), 500 fmol of each discrimination primer and 750 fmol of common primer (Fig. 3C). Sets of discrimination and common primers were synthesized to perform LDR on the expected conversion products and varied at the bases (B_i) corresponding to the *Msp*I position sense strand ($B_1B_2B_3B_4$ = CCGG for wild-type). The discrimination primers had wild-type sequence and terminate in -B_1B_2(-OH-3′). The discrimination primers were synthesized as a set of four primers each with C, T, G and A in turn at B_2. The common LDR primers had (5′-PO_4-)B_3B_4- followed by wild-type sequence and hybridized to the template with its 5′ base adjacent to the 3′ base of a discrimination primer. Discrimination

primers varied the 3′-terminal base to identify error products at B_2 of the *Msp*I position. For simplicity, only B_2 was monitored. LDR primers matched the expected conversion products; for example, conversion of -CCGG- template to -ACGT- required discrimination primers ending in -AC, -AT, -AG and -AA and a common primer with 5′-pGT-. Discrimination primers had 5′ tails of different length and a FAM label for fluorescence detection. The tail length allowed physical separation of different LDR products on an acrylamide gel and thus identification of the LDR products.

LDR reactions were preincubated for 1.5 min at 94°C prior to the addition of 5 nmol *Tth* ligase, followed by 10 LDR cycles of 94°C for 15 s, 65°C for 2 min and a final hold briefly at 94°C. Reactions were cold quenched and stored at –70°C. The LDR products were separated on 10% acrylamide gels containing 7 M urea, with 0.6× TBE (1× TBE is 90 mM Tris base, 90 mM borate, 2 mM EDTA) in the gel and running buffer. Data were collected using an ABI 373 DNA sequencer with Genescan 672 software.

Image processing

Gel pictures were produced by the ABI 672 Analysis software. Dye-specific images were opened in Adobe Photoshop 3.0, cropped, resized and converted to grayscale. The grayscale images were opened in NIH Image 1.59, inverted and 1D vertical background was subtracted. The background subtracted images were reinverted and rendered in pseudocolor by Photoshop to make intensity differences easier to compare. Except for color replacement, only linear image processing was performed to preserve relative intensities.

RESULTS AND DISCUSSION

Initial experiments were designed to determine the efficiency of generating PCR products when using primers containing 3′-terminal nucleotide analogs (Materials and Methods). Eight different analogs were designed to pair with more than one of the four natural bases in order to convert one base to another base at a specific position in a sequence. Primer pairs containing either a nucleotide analog or one of the four natural bases at their 3′-ends were used to amplify four different templates (Fig. 3A). Each nucleotide analog and natural base was mispaired (or paired) in turn with all four natural bases on the opposite strand and amplification was attempted with either *Taq* Stoeffel Fragment or *Taq* Fluorescent Sequencing polymerases. The relative amplification efficiency was determined by the number of cycles required to generate visible product on an ethidium bromide stained agarose gel (Table 1). We found that both *Taq* Stoeffel Fragment and *Taq* Fluorescent Sequencing polymerases produced comparable amounts of product (data not shown). Perfectly matched natural base primers generated visible product after 10 cycles, however, some analog primers generated no product after 50 cycles. The analogs that did amplify with high efficiency were those that were best able to 'read' the opposite strand sequence (Fig. 1).

One product for each analog (as well as the natural base controls) was reamplifed and sequenced to determine polymerase preference in inserting nucleotide bases opposite the analog (Table 1). We found that the Q_1, Q_5, Q_6, Q_{16} and Q_{18} primers generated detectable true conversion product, however, only Q_5 primers generated almost exclusively true conversion product. No single analog functioned as a 'universal base' (26) capable of generalized conversion. Unexpectedly, some products contained sequences that were difficult to read across the middle four bases, suggesting single base insertions or deletions occurred during PCR extension. This was especially prevalent in products generated from mismatched natural bases (see below).

Table 1. Extension efficiency and conversion with 3′ natural base and nucleotide analog primers

Primer 3′ base	TCGA template reads A writes (efficiency)	CCGG template reads G writes (efficiency)	GCGC template reads C writes (efficiency)	ACGT template reads T writes (efficiency)
T	A (+++)	A (++)	(++)	(++)
C	(++)	G (+++)	(++)	(++)
G	(++)	(++)	C (+++)	(+++)
A	(+)	(+)	T (+++)	T (+++)
Q_1	A,T (±)	(±)[a]	(–)	(+++)[a]
Q_2	(±)[a]	(±)[a]	(–)	T (++)[a]
Q_5	(++)	(++)	(+++)	C (+++)
Q_6	A,G (+++)	(+++)	(++)	(++)
Q_7	(+)	(+)	(+++)	T (+++)
Q_{16}	A,T (+)[a]	(–)	(–)	(–)
Q_{18}	(+)[a]	(±)[a]	(±)[a]	T,A (+++)
Q_{19}	A (++)[a]	(–)	(±)[a]	(+)[a]

[a]Low product yield.

Four different templates were used to test primer extension from a 3′ base or analog paired in turn with A, G, C and T. Relative efficiency was determined by the number of cycles required to generate visible product with *Taq* Stoeffel Fragment polymerase: (+++), 10 cycles; (++), 20 cycles; (+), 30 cycles; (±), 40–50 cycles; (–), no product. Two of the natural base mismatch primer products were sequenced. Generally, the most efficiently amplified template for each analog was reamplified with truncated primers and sequenced to determine which bases are written opposite each analog. In one case (Q_1) a lower efficiency extension product with higher yield was selected for sequencing. Mixed base writing preference for some analogs is indicated, with most frequent product listed first.

To test the ability of convertides to reduce mismatch extension errors, we assessed the effects of preconversion PCR cycles on fidelity. PCR products generated from template amplified with only natural base conversion primers were compared to products resulting from two initial PCR cycles using convertides followed by selective amplification using specific natural base primers. We performed preconversion PCR with primer pairs containing Q_5, Q_6 and Q_7 analogs, since these convertides had been shown to be the most efficiently extended. To improve overall PCR fidelity and 3′ mismatch primer extension, CiNF buffer (Materials and Methods) was used (33). Nine different synthetic duplex templates containing mutated *Msp*I sites were amplified with or without preconversion using 3′ analog preconversion primers. Both natural base conversion primers and 3′ analog preconversion primers were designed to manipulate the outside bases CCGG of the *Msp*I position (Fig. 3). Some conversions were intended to serve as controls. In these cases, the original bases in the template were either restored after analog preconversion or never changed with full-length perfect match primers. All steps were performed identically between preconversion and non-preconversion reactions, except that preconversion reactions used as template the product of two cycles of convertide PCR for succeeding rounds of amplification,

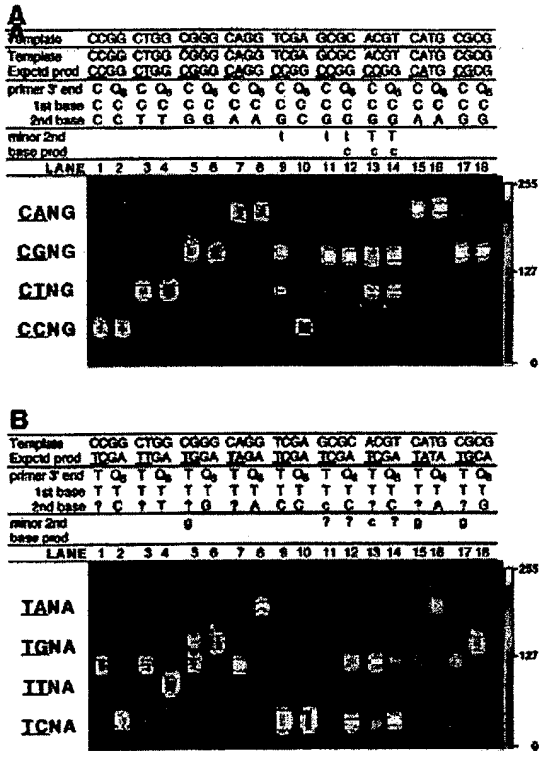


Figure 4. Conversion by natural base and Q_6 convertide. Conversion products from nine templates were detected by PCR/LDR (Materials and Methods). Each template was a 50 bp synthetic duplex DNA of identical sequence except for the central four bases which have the sequence indicated. Conversion occurred within these four bases. The expected conversion products produced by starting with the conversion primers having the indicated 3′ natural base or convertide are shown. (A) Conversion of the first base to C with and without Q_6 preconversion. (B) Conversion of the first base to T with and without Q_6 preconversion.

while synthetic duplex served as the starting material for PCRs with no preconversion. In both cases, 3′ natural base primers were used to selectively amplify the desired end product. These primers contained non-hybridizing zipcode sequences on their 5′-ends (Materials and Methods), which ultimately served as primer binding sites for the final 20–30 cycles of PCR (Fig. 3B). Conversion products were quantified by LDR (Fig. 3C).

We found that overall, natural base mismatch conversion generated >80% incorrect conversion products (Fig. 4A, lane 9, and B, lanes 1, 3, 5, 7, 15 and 17), but preconversion could improve the fidelity and/or the yield of some conversions. In general, transversions were difficult to achieve even with preconversion. G→C and A→C conversion generated very little of the expected product with either the natural base or Q_6 primers (Fig. 4A, lanes 11–14). Use of Q_6 preconversion improved the yield of G→T and A→T conversion products (compare natural base conversion in Fig. 4B, lanes 11 and 13, with Q_6 preconversion in lanes 12 and 14). In the case of transitions, C→T conversion produced unexpected one base shortened artifacts with natural base mismatch primers on the CXGG templates (Fig. 4B, lanes 1, 3, 5, 7, 15 and 17), but the correct products were generated when using Q_6 preconversion (Fig. 4B, lanes 2, 4, 6, 8, 16 and 18). In addition, Q_6 primers did improve the yield of the expected T→C conversion product (Fig. 4A, lanes 9 and 10). The controls performed as expected: all C→C and T→T non-conversion reactions worked correctly without convertides (Fig. 4A, lanes 1, 3, 5, 7, 15 and 17, and B, lane 9) and the corresponding Q_6 preconversion products were restored to the original sequence (Fig. 4A, lanes 2, 4, 6, 8, 16 and 18, and B, lane 10). In summary, Q_6 preconversion reduced or eliminated artifacts produced by natural base C→T and T→C conversion and facilitated transitions in general. Transversions were only partially successful: G→T and A→T conversions could be improved with preconversion, but G→C and A→C conversion could not be achieved.

Apparently correct conversions were observed with attempted C→G and C→A transversions, however, carefully designed control templates revealed that these 'conversions' were artifactual. C→G and C→A conversion appeared to be successful for templates containing a central CpG dinucleotide (Fig. 5A and B, lanes 1–3 and 13–21). However, the same final conversion products were observed with other templates lacking the central CpG dinucleotide, now clearly incorrect. For example, a GCGC product resulted during G conversion in reactions where the expected product should have contained T, G or A in the second position (Fig. 5A, lanes 4–12). Also, an ACGT product resulted during A conversion where the expected product should have inserted a non-C base in the second position (Fig. 5B, lanes 4–12 and 22–27). The mismatch primers used to alter the outer bases of the recognition site did not reach the central dinucleotide, yet these bases were altered. It is doubtful the 'successful' conversions occurred through the intended mechanism and thus represent fortuitous artifacts. The yield of LDR product was low for two palindromic templates despite efficient PCR (Fig. 5A and B, lanes 22–27). These conversion reaction products presumably contain a large fraction of insertions or deletions, which cannot be detected by the current set of LDR primers. In summary, C→G conversion was partially accomplished by both Q_5 (Fig. 5A, lanes 5, 8, 11 and 23) and the natural base G (Fig. 5A, lanes 4, 7, 10 and 22), however, preconversion does not appear to improve conversion. C→G conversion exhibits sequence dependence.

The results of the preconversion study indicate that errors in natural base conversion were prevalent, but the use of Q_5, Q_6 and Q_7 convertides in preconversion reduced polymerase error in certain cases. In terms of conversion reactions, transitions were easier to accomplish than transversions. This is in agreement with previous findings. Newton *et al.* observed more errors in extension of primers with 3′-terminal C·T, A·A and T·T mismatches (transversions) than with purine-pyrimidine mismatches (transitions) (34). In our hands, pyrimidine–pyrimidine conversion usually generated the expected product, especially when using convertides. In cases of purine–pyrimidine and pyrimidine–purine conversion, incorrect products were often generated (summary of results in Table 2). Formation of incorrect conversion products can be explained in part by a transient base pair slippage of the primer 3′ nucleotide (or analog) to a misaligned position on the template (Fig. 6). As a result, the sequence following the mismatch is not complementary to the original template. Consistent with this hypothesis is the observation of unreadable sequence immediately following the analog in our initial sequencing experiments. Palindromic products, especially

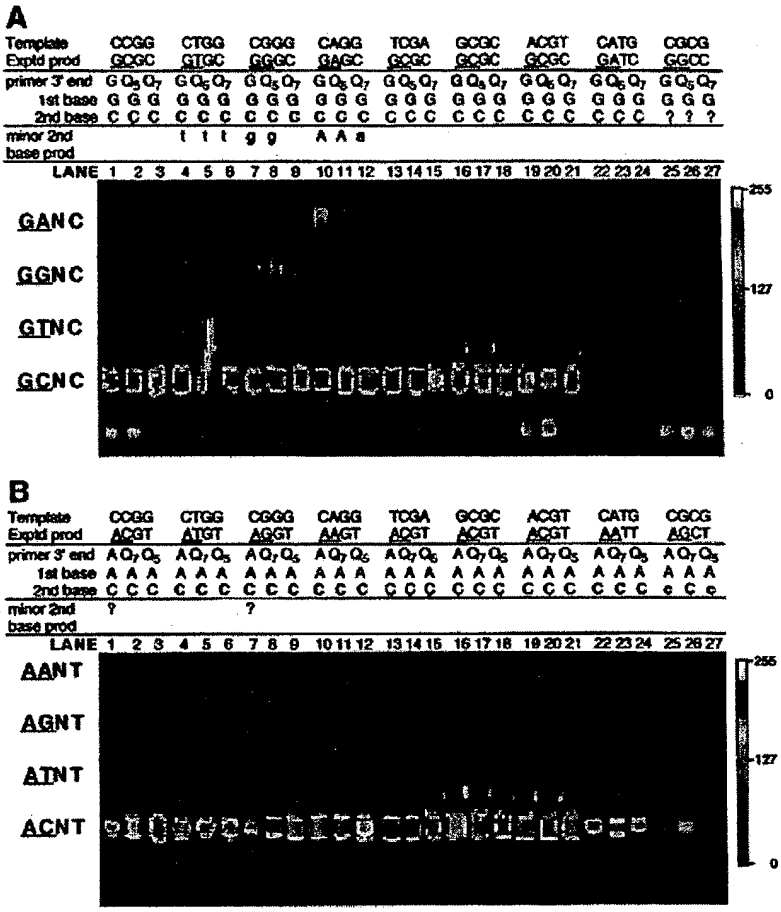


Figure 5. Conversion by natural base and Q_5 and Q_7 convertides. Conversion products from nine templates were detected by PCR/LDR (Materials and Methods). Each template was a 50 bp synthetic duplex DNA of identical sequence except for the central four bases which have the sequence indicated. Conversion occurred within these four bases. The expected conversion products produced by starting with the conversion primers having the indicated 3′ natural base or convertide are shown. (**A**) Conversion of the first base to G with and without Q_5 or Q_7 preconversion. (**B**) Conversion of the first base to A with and without Q_5 or Q_7 preconversion.

CpG dinucleotides, are themselves prone to slippage and extension. We observed palindromic products were frequently produced from non-palindromic templates. These artifacts were reduced by the presence of 10% formamide in the PCR buffer, presumably through destabilization of misaligned structures. Finally, nucleotide analogs produced fewer artifacts than natural bases. Different analogs produced different kinds and quantities of artifacts, perhaps according to their relative ability to base pair and stabilize a slippage misalignment. Thus, if polymerase extension is slow from an analog poorly base paired with the template, extension from a strong transient base pair generated by slippage could exceed the rate of extension from a weakly base paired 3′-terminal base.

As discussed earlier, PCR–RFLP has been widely used to detect rare mutations. A limitation of this technique is reliance on serendipity to yield mutations that can be modified to create restriction sites in either the wild-type or the mutant template. A second limitation imposed on this approach is the need to avoid using 3′-terminal mismatch primers, since extension from these primers is known to be error prone. To date, the majority of successful attempts have used interrupted palindromic restriction sites to avoid using 3′-terminal mismatch primers. Mutations in the cancer-causing genes K-*ras* and H-*ras* were detected at a sensitivity of 1 in 10^5 using PCR–RFLP with interrupted palindromic enzymes *Xmn*I (9), *Alw*NI (35) and *Bst*NI or *Mva*I (36,37). These PCR–RFLP experiments and others (18,38–42) avoid 3′-terminal mismatches, however, most cancer mutations are in sequences that cannot be converted to interrupted palindromes, for example at CpG dinucleotides.

Table 2. Most effective conversion (Figs 4 and 5)

Starting template[a]	First base converted to			
	C	T	G	A
1 CCGG	C	Q_6	Q_7 (FP)	Q_5 (FP)
2 CTGG	C	Q_6	X (err C)	X (err C)
3 CGGG	C	Q_6	Q_5 (err C)	X (err C)
4 CAGG	C	Q_6	G (err C)	X (err C)
5 TCGA	Q_6	T or Q_6	Q_7 (FP)	Q_5 (FP)
6 GCGC	X (err G)	Q_6	G	Q_5 or Q_7
7 ACGT	X (err G)	Q_6 weak	Q_7	A or Q_7
8 CATG	C	X	Q_5 (err C)	X (err C)
9 CGCG	C	Q_6	X	Q_7 (err C)

[a]The 50 bp synthetic duplex DNA templates containing p53 sequence spanning codon 248 are distinguished by the four bases replacing the *Msp*I site, which are shown.

Nine duplex DNA templates were used in conversion reactions. Each contained sequence identical to p53 surrounding codon 248, except the *Msp*I site was replaced by a different four base sequence ($B_1B_2B_3B_4$). B_1 and B_4– (opposite strand) were simultaneously converted in turn to C, T, G and A either directly by PCR using natural base primers or by preconversion PCR with nucleotide analog primers followed by PCR with natural base primers. In non-conversion control reactions the 'conversion' product is identical to the original template. A natural base is used to indicate control reactions and cases in which preconversion did not improve conversion. Preconversion was performed using Q_6 to facilitate conversion to C and T and using Q_5 and Q_7 to facilitate conversions to G and A. Conversion primers determine B_1 and B_4; LDR was performed to detect unintended base changes in B_2 (which ideally is unchanged after conversion). Conversion improved by preconversion is indicated by the nucleotide analog used. Preconversion equally as effective in control reactions as natural base primers is also indicated by the analog used. Low conversion fidelity results in large B_2 error. Major B_2 error products are identified (e.g. err C indicates C at B_2) and the absence of correct product indicated no conversion method was successful (X, no correct product). Apparently correct product probably formed through a fortuitous mechanism is indicated (FP, false positive).

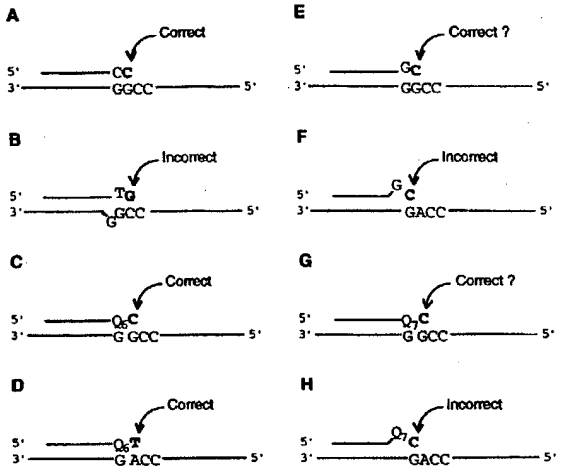


Figure 6. Fidelity of polymerase extension. Primer slippage accounts for many of the observed products of extension (Figs 4 and 5). (**A**) Perfectly complimentary primer gives correct product. (**B**) T:G mismatch at the second base explains TGGA (or TGCA) product. (**C**) Extension from a Q_6:G pairing with no slippage on the minus strand of the CCGG template (followed by 3′ T conversion primers) resulted in the expected TCGA product. (**D**) Extension from a Q_6:G pairing with no slippage on the minus strand of the CTGG template and several other templates (followed by 3′ T conversion primers) resulted in the expected products. (**E**) G·G mismatch extension apparently gave the expected GC product on one template, but perhaps only fortuitously (see F). (**F**) All extensions from G·G mismatches gave GC extension products, consistent with a G·T mismatch formed by slippage at the preceding base (Fig. 3). (**G**) Q_5:G and Q_7:G extension products apparently gave the expected GC product on one template, but perhaps only fortuitously (see H). (**H**) All extensions from Q_5:G and Q_7:G mismatches (followed by 3′ G conversion primers) gave GC extension products consistent with a Q_5:T or Q_7:T mismatch at the preceding base (Fig. 3).

A larger fraction of mutations would be made into targets for detection if contiguous recognition sequences could be introduced with as few errors as interrupted palindromic recognition sequences. Currently, contiguous restriction sites are introduced by terminal 3′ mismatch primer extension, which is prone to errors. O'Dell *et al.* tested a general method for introducing different restriction sites at CpG dinucleotides using mismatch PCR (19). The outer bases of four different CpG dinucleotides in the human LDL receptor gene were altered to create *Taq*I (TCGA), *Msp*I (CCGG) or *Hha*I (GCGC) sites. In these targets, *Taq*I sites were successfully generated by 3′ T mismatch primers. The method was able to detect homozygous and heterozygous individuals, however, the ratio of products representing each allele was not equal, as is expected in germline mutations. We have shown several cases where T mismatch conversion failed to create a *Taq*I site, thus the method is sequence dependent. O'Dell *et al.* found that C and G mismatch conversion failed. We agree with their conclusion that stronger base pairing leads to mispriming, possibly through stabilization of primer slippage on the template. Gotoda *et al.* claim to have successfully used PCR–RFLP to introduce an *Mae*II site (ACGT) by extension of a 3′ C·A mismatch to produce a T→C transition (43). Athma *et al.* used PCR extension of a 3′-terminal mismatch primer to create a restriction site for discriminating between two alleles (44). A G·T mismatch produced a *Mva*I site (CC A/T GG) through an A→G transition. We successfully performed A→G conversion using a natural base mismatch, but encountered difficulties with T→C conversion by natural base primers. In our hands, transitions can be accomplished more readily than transversions, but the yield of correct product can be sequence dependent. Others have also found that PCR–RFLP can produce false positive results (20). Our use of the ligase detection reaction allowed us to determine the precise amounts of misextension products generated.

We have measured the fidelity of polymerase extension from primers containing 3′ natural bases and nucleotide analogs. Our results indicate that natural base mismatch primer extension cannot be used as a general technique to create restriction sites in any given sequence for RFLP analysis. Primer slippage appears to be an important mechanism for producing error in mismatch primer extension. This source of error may have a dramatic impact on some allele-specific PCR and other methods of high sensitivity mutation detection. With further development and testing of nucleotide analogs to facilitate conversion, mismatch primer extension may become a technique that can efficiently introduce desired mutations with few artifacts. We have found some nucleotide analogs improve mismatch primer extension (Table 3). Further improvement of 3′ mismatch extension will be required to minimize the high degree of context-dependent error

seen in transversions and lead to improved levels of sensitivity and broader scope of PCR–RFLP-based mutation detection.

Table 3. Summary of conversion strategy

Starting base	Conversion to			
	C	T	G	A
C	C	Q_6		
T	Q_6	T		
G			G	A or Q_7
A			Q_5 or Q_7	A

A Q_n convertide indicates preconversion is required using the indicated convertide prior to final conversion using natural base primers. In some cases, an additional convertide or using only the natural base will result in the desired conversion.

ACKNOWLEDGEMENTS

We thank Michael Wigler, Thierry Soussi, Mark Sobel, Jerard Hurwitz, Saul Silverstein, Harry Ostrer, Michael Osborne, Daniel Knowles, Pat Paty, Reyna Favis and members of the Barany and Paty laboratories for technical assistance and helpful discussion. This work was supported by grants from the National Cancer Institute (CA65930-02), Strang Cancer Prevention Center and the Applied Biosystems Division of the Perkin Elmer Corporation. We also thank Peiming Zhang and Travis Johnson from the Bergstrom laboratory who synthesized the Q_1, Q_2 and Q_{18} primers and Melissa Cameron from the Hammer laboratory who synthesized the Q_{16} and Q_{19} primers.

REFERENCES

1 Saiki,R.K., Bugawan,T.L., Horn,G.T., Mullis,K.B. and Erlich,H.A. (1986) *Nature*, **324**, 163–166.
2 Sidransky,D., Tokino,T., Hamilton,S.R., Kinzler,K.W., Levin,B., Frost,P. and Vogelstein,B. (1992) *Science*, **256**, 102–105.
3 Brennan,J.A., Mao,L., Hruban,R.H., Boyle,J.O., Eby,Y.J., Koch,W.M., Goodman,S.N. and Sidransky,D. (1995) *N. Engl. J. Med.*, **332**, 429–435.
4 Cha,R.S., Zarbl,H., Keohavong,P. and Thilly,W.G. (1992) *PCR Methods Applicat.*, **2**, 14–20.
5 Haliassos,A., Chomel,J.C., Grandjouan,S., Kruh,J., Kaplan,J.C. and Kitzis,A. (1989) *Nucleic Acids Res.*, **17**, 8093–8099.
6 Chen,Z.-Y. and Zarbl,H. (1997) *Anal. Biochem.*, **244**, 191–194.
7 Sandy,M.S., Chiocca,S.M. and Cerutti,P.A. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 890–894.
8 Pourzand,C. and Cerutti,P. (1993) *Mutat. Res.*, **288**, 113–121.
9 Kumar,R. and Barbacid,M. (1988) *Oncogene*, **3**, 647–651.
10 Kumar,R. and Dunn,L.L. (1989) *Oncogene Res.*, **4**, 235–241.
11 Jacobson,D.R. and Mills,N.E. (1994) *Oncogene*, **9**, 553–563.
12 Jacobson,D.R. and Moskovits,T. (1992) *PCR Methods Applicat.*, **1**, 299.
13 Chen,J. and Viola,M.V. (1991) *Anal. Biochem.*, **195**, 51–56.
14 Di Giuseppe,J.A., Hruban,R.H., Offerhaus,G.J., Clement,M.J., van den Berg,F.M., Cameron,J.L. and van Mansfeld,A.D. (1994) *Am. J. Pathol.*, **144**, 889–895.
15 Kahn,S.M., Jiang,W., Culbertson,T.A., Weinstein,B., Williams,G.M., Tomita,N. and Ronai,Z. (1991) *Oncogene*, **6**, 1079–1083.
16 Levi,S., Urbano-Ispizua,A., Gill,R., Thomas,D.M., Gilbertson,J., Foster,C. and Marshall,C.J. (1991) *Cancer Res.*, **51**, 3497–3502.
17 Mitsudomi,T., Viallet,J., Mulshine,J.L., Linnoila,R.I., Minna,J.D. and Gazdar,A.F. (1991) *Oncogene*, **6**, 1353–1362.
18 Hattori,M., Nanko,S., Dai,X.Y., Fukuda,R. and Kazamatsuri,H. (1994) *Biochem. Biophys. Res. Commun.*, **202**, 757–763.
19 O'Dell,S.D., Humphries,S.E. and Day,I.N. (1996) *Genome Res.*, **6**, 558–568.
20 Hodanova,K., Hrebicek,M., Cervenkova,M., Aerts,J.M. and Zeman,J. (1997) *J. Inherit. Metab. Dis.*, **20**, 611–612.
21 Hoops,G.C., Zhang,P., Johnson,W.T., Paul,N., Bergstrom,D.E. and Davisson,V.J. (1997) *Nucleic Acids Res.*, **25**, 4866–4871.
22 Brown,D.M. and Lin,P.K.T. (1991) *Carbohydrate Res.*, **216**, 129–139.
23 Bergstrom,D.E., Zhang,P. and Johnson,W.T. (1996) *Nucleosides Nucleotides*, **15**, 59–68.
24 Bergstrom,D.E., Zhang,P., Toma,P.H., Andrews,P.C. and Nichols,R. (1995) *J. Am. Chem. Soc.*, **117**, 1201–1209.
25 Zhang,P., Johnson,W.T., Klewer,D., Paul,N., Hoops,G., Davisson,V.J. and Bergstrom,D.E. (1998) *Nucleic Acids Res.*, **26**, 2208–2215.
26 Hill,F., Loakes,D. and Brown,D.M. (1998) *Proc. Natl Acad. Sci. USA*, **95**, 4258–4263.
27 Barany,F. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 189–193.
28 Barany,F. (1991) *PCR Methods Applicat.*, **1**, 5–16.
29 Day,D.J., Speiser,P.W., White,P.C. and Barany,F. (1995) *Genomics*, **29**, 152–162.
30 Khanna,M., Park,P., Zirvi,M., Cao,W., Picon,A., Day,J., Paty,P. and Barany,F. (1999) *Oncogene*, **18**, 27–38.
31 Johnson,W.T., Zhang,P. and Bergstrom,D.E. (1997) *Nucleic Acids Res.*, **25**, 559–567.
32 Pon,R.T., Usman,N. and Ogilvie,K.K. (1988) *BioTechniques*, **6**, 768–775.
33 Day,J.P., Berstrom,D., Hammer,R.P. and Barany,F. (1999) *Nucleic Acids Res.*, **8**, 1819–1827.
34 Newton,C.R., Graham,A., Heptinstall,L.E., Powell,S.J., Summers,C., Kalsheker,N., Smith,J.C. and Markham,A.F. (1989) *Nucleic Acids Res.*, **17**, 2503–2516.
35 Anderson,J.A., Irish,J.C. and Ngan,B.Y. (1992) *J. Otolaryngol.*, **21**, 321–326.
36 Urban,T., Ricci,S., Grange,J.D., Lacave,R., Boudghene,F., Breittmayer,F., Languille,O., Roland,J. and Bernaudin,J.F. (1993) *J. Natl Cancer Inst.*, **85**, 2008–2012.
37 Ronai,Z. and Minamoto,T. (1997) *Hum. Mutat.*, **10**, 322–325.
38 Beutler,E., Gelbart,T. and West,C. (1990) *Clin. Chim. Acta*, **194**, 161–166.
39 Hingorani,A.D. and Brown,M.J. (1995) *Biochem. Biophys. Res. Commun.*, **213**, 725–729.
40 Kuwata,S., Yanagisawa,M., Saeki,H., Nakagawa,H., Etoh,T., Tokunaga,K., Juji,T. and Shibata,Y. (1995) *J. Allergy Clin. Immunol.*, **96**, 1051–1060.
41 Nishiwaki,Y., Kamino,K., Yoshiiwa,A., Nagano,K., Takeda,M., Tanabe,H., Nishimura,T., Kobayashi,T., Yamamoto,H., Nonomura,Y., Yoneda,H., Sakai,T., Imagawa,M., Miki,T. and Ogihara,T. (1996) *Clin. Genet.*, **49**, 119–123.
42 Ishihara,M., Ohno,S., Ishida,T., Naruse,T., Kagiya,M., Mizuki,N., Maruya,E., Saji,H. and Inoko,H. (1997) *Tissue Antigens*, **49**, 107–110.
43 Gotoda,T., Yamada,N., Murase,T., Shimano,H., Shimada,M., Harada,K., Kawamura,M., Kozaki,K. and Yazaki,Y. (1992) *J. Lipid Res.*, **33**, 1067–1072.
44 Athma,P., Fidahusein,N. and Swift,M. (1995) *Biochem. Biophys. Res. Commun.*, **210**, 982–986.

EXHIBIT 5

Nucleic Acids Research, 1999, Vol. 27, No. 8 1819–1827

Nucleotide analogs and new buffers improve a generalized method to enrich for low abundance mutations

Joseph P. Day, Robert P. Hammer[1], Donald Bergstrom[2] and Francis Barany*

Department of Microbiology, Box 62, Hearst Microbiology Research Center, Strang Cancer Prevention Center, Joan and Sanford I. Weill Medical College of Cornell University, 1300 York Avenue, New York, NY 10021, USA, [1]Department of Chemistry, Louisiana State University, Baton Rouge, LA 70803, USA and [2]Department of Medicinal Chemistry and Molecular Pharmacology, Purdue University, West Lafayette, IN 46208, USA

Received as resubmission January 19, 1999; Revised and Accepted February 18, 1999

ABSTRACT

A high sensitivity method for detecting low level mutations is under development. A PCR reaction is performed in which a restriction site is introduced in wild-type DNA by alteration of specific bases. Digestion of wild-type DNA by the cognate restriction endonuclease (RE) enriches for products with mutations within the recognition site. After reamplification, mutations are identified by a ligation detection reaction (LDR). This PCR/RE/LDR assay was initially used to detect PCR error in known wild-type samples. PCR error was measured in low $|\Delta pK_a|$ buffers containing tricine, EPPS and citrate, as well as otherwise identical buffers containing Tris. PCR conditions were optimized to minimize PCR error using perfect match primers at the *Msp*I site in the p53 tumor suppressor gene at codon 248. However, since mutations do not always occur within pre-existing restriction sites, a generalized PCR/RE/LDR method requires the introduction of a new restriction site. In principle, PCR with mismatch primers can alter specific bases in a sequence and generate a new restriction site. However, extension from 3′ mismatch primers may generate misextension products. We tested conversion of the *Msp*I (CCGG) site to a *Taq*I site (TCGA). Conversion was unsuccessful using a natural base T mismatch primer set. Conversion was successful when modified primers containing the 6*H*,8*H*-3,4-dihydropyrimido[4,5-*c*][1,2]oxazine-7-one (Q_6) base at 3′-ends were used in three cycles of preconversion PCR prior to conversion PCR using the 3′ natural base T primers. The ability of the pyrimidine analog Q_6 to access both a T-like and C-like tautomer appears to greatly facilitate the conversion.

INTRODUCTION

High sensitivity medical diagnostic assays depend on accurate DNA amplification by DNA polymerases (1–4). Such DNA-based diagnostic methods are needed, for example, to improve cancer staging and aid clinical decisions through molecular characterization of the disease. Low level mutations may be detected by cloning PCR-amplified fragments and accurately quantified by probing for mutant DNA using allele-specific oligonucleotides (ASOs), however, this process requires several days to complete (5,6). Alternatively, allele-specific PCR (AS-PCR) can be used to detect low abundance mutations. By designing primers with one or more mismatches, mutant DNA template can be efficiently extended, while poor extension is achieved on wild-type DNA template. However, once these primers extend with or without a mismatch, the products thereafter are perfect matches for the primer in subsequent PCR cycles. Thus, false positive signals are amplified in subsequent cycles. Moreover, PCR error can generate a base change in the template which perfectly matches the primer. AS-PCR can detect pyrimidine↔purine transversions at sensitivities of 1 in 10^5 (7,8). Nevertheless, the majority of cancer-associated mutations are C↔T and A↔G transitions, for example, >80% of p53 point mutations (9). A DNA diagnostic method is required to accurately quantify this type of low abundance mutation.

The ligation detection reaction (LDR) uses two adjacent primers and a thermostable ligase to distinguish all four bases potentially found at any position in a DNA sequence (10–13). Thermostable ligase demonstrates the highest fidelity when the discriminating base is located at the 3′-end of the upstream primer (14). PCR/LDR (PCR of a sequence from genomic DNA followed by LDR) can detect mutations with a sensitivity of approximately one mutant allele in 4000 normal alleles (13). Sensitivity of approximately 1 in 10^6 has been achieved by combining PCR with restriction endonuclease (RE) digestion of wild-type DNA (15,16). Mutations occurring within the restriction site prevent cleavage of the mutant allele, while wild-type alleles bearing canonical restriction site sequence are depleted. As a result, subsequent PCR cycles preferentially amplify mutant DNA. If a mutation site is not within an endonuclease recognition site present in wild-type DNA, a restriction site must be introduced. This is typically done by PCR using a primer or primers with mismatched bases. Mutations cannot be detected in any portion of the restriction site spanned by the primers, since those bases are introduced directly through the primers. In a random DNA sequence, >20% of bases are contained within a pre-existing four base restriction site and 60% of bases are within a four base subsequence that can be converted into a restriction site by a single base change. In these small sites, 3′-terminal base mismatch primers must frequently be used. While conceptually

*To whom correspondence should be addressed. Tel: +1 212 746 4288; Fax: +1 212 746 8588; Email: barany@mail.med.cornell.edu

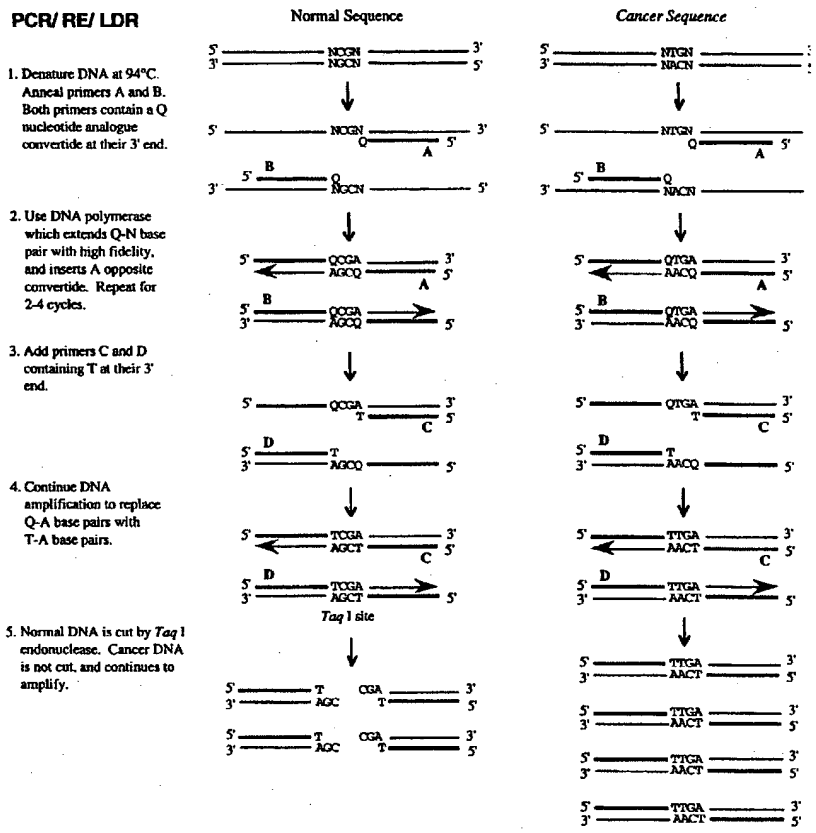


Figure 1. Preferential amplification of mutant DNA by PCR/RE/LDR. Preconversion (steps 1 and 2) using a nucleotide analog (Q) in mutant and wild-type DNA allows more efficient introduction of a restriction site at a CpG dinucleotide. The addition of natural base primers (step 3) completes conversion, replacing the nucleotide analog with the desired normal base. Finally, wild-type DNA is digested with the RE for which the site was created (step 5). Mutant DNA is subsequently reamplified.

straightforward, 3′ mismatch extension has proven to be difficult (7,17–19). The introduction of interrupted palindromic restriction sites has been more successful using internal mismatch primers spanning one half-site through the intervening bases up to the other half-site (20,21). Several perfectly matched bases stabilize the 3′-end of the mismatch primer. This approach may be used only if the second half-site is present naturally in wild-type DNA.

REs recognizing interrupted palindromes are less abundant than endonucleases recognizing contiguous four and six base sites. Multiple base changes would often be required to introduce an interrupted palindrome restriction site to identify mutations at any base. In order to develop a general approach, this work introduces contiguous four base Type II restriction sites in wild-type sequences containing a central CpG dinucleotide by altering one base on each side of the CpG. CpG dinucleotides are frequent sites of mutation; for example, ~40% of the mutations observed in the p53 tumor suppressor gene fall into this category (9). We tested conversion of the *MspI* site (CCGG) to a *TaqI* site (TCGA) in p53 exon 7 at codon 248 by PCR using a set of 3′ mismatch primers which alter the outer bases of the *MspI* site (Fig. 1). Preconversion using 3′ nucleotide analog (Q) primers may be performed prior to adding natural base primers to avoid extension from primers with 3′ natural base mismatches. Conversion was performed with and without preconversion to determine whether preconversion facilitates conversion. This process introduces a restriction site that allows digestion of wild-type DNA with an RE and leads to preferential amplification of the undigested mutant DNA. Mutant products are quantified by LDR. This PCR/RE/LDR assay is a high sensitivity variant of PCR/LDR.

Techniques, such as PCR/RE/LDR, that rely on mutant enrichment require optimization of reaction conditions in order to minimize PCR errors. These errors would be indistinguishable from mutations originally present in clinical samples. Standard PCR buffers contain Tris, however its pK_a is strongly dependent on temperature. A PCR reaction containing Tris pH 8.3 (measured at 23°C) is ~pH 7 near 65°C (the extension temperature) and drops to ~pH 6 near 95°C (the template melting temperature). PCR error can result from template degradation and polymerase misincorporation. Template degradation occurs during periods of high temperature and low pH in each PCR cycle and limits product size in 'long' PCR (22–24). Raising the buffer pH in long PCR (using Tris 9.1) reduces the amount of template cleavage and increases PCR efficiency (22). Although the efficiency of long PCR increases with higher pH, the level of mutations within these PCR products may also increase since high pH decreases the fidelity of *Taq* and *Pfu* polymerases (25–27). Use of alternative PCR buffers

with lower $|\Delta pK_a|$ can improve polymerase fidelity and still reduce template damage by maintaining more neutral pH over a wider temperature range (26,28). The addition of glycerol or formamide may reduce mutations arising from template damage during PCR cycling and may help avoid misextension from mispaired primers (2,29).

We tested proofreading and non-proofreading thermostable DNA polymerases in several PCR buffers formulated on the basis of an analysis of known sources of PCR error. Our test PCR buffers contained low $|\Delta pK_a|$ buffering compounds and different salts. We show that mismatch extension is prone to error far beyond that produced by polymerase error or template degradation during PCR. Directly probing PCR/RE products using LDR allows identification of specific mutations and quantification of each mutation produced. PCR fidelities using buffers with and without formamide were compared. The use of low $|\Delta pK_a|$ buffers with formamide greatly reduces background PCR error. Preconversion with 3′ nucleotide analog primers significantly improved the fidelity of base conversion to introduce a new restriction site.

MATERIALS AND METHODS

Oligonucleotide synthesis

Oligonucleotides were synthesized at the 0.2 µmol scale by cyanoethyl phosphoramidite chemistry on an Applied Biosystems 394 DNA synthesizer. Standard 500 Å CPG columns and reagents (Applied Biosystems) were used with the following exceptions. Oligonucleotides 50 bases in length were synthesized using wide pore 1000 Å CPG columns (Applied Biosystems). Oligonucleotides with fluorescent dye FAM at the 5′-terminus were synthesized using FAM phosphoramidite (Applied Biosystems) with a 15 min coupling step. Oligonucleotides with 5′ phosphate were synthesized using phosphorylation reagent (Glen Research) with a 15 min coupling step. Oligonucleotides with 3′ blocking group were synthesized using 3′-Spacer CPG columns (Glen Research). Oligonucleotides with the 3′ nucleotide analog 6-(2′-deoxy-β-D-ribofuranosyl)-6*H*,8*H*-3,4-dihydro-pyrimido-[4,5-*c*]-[1,2]oxazine-7-one (Q_6) were synthesized using dP-CPG (Glen Research).

PCR polymerases and buffers

The polymerases used were Ampli*Taq* (Applied Biosystems), Vent and Vent(exo–) (New England Biolabs) and Expand polymerase mix (*Taq* and *Pwo*I polymerase mixture, in the Expand High Fidelity kit; Boehringer Mannheim). The commercially available PCR buffers used were supplied in the Ampli*Taq* and Expand kits. Tris pH 9.1 (pH values were measured using 1 M stock solutions at 23°C), tricine pH 8.7, EPPS [*N*-(2-hydroxyethyl)piperazine-*N*′-3-propanesulfonic acid] pH 8.4 and citrate pH 7.6 (Sigma) were used for alternative PCR buffers. Unless otherwise noted, each 20 µl reaction contained 20 mM Tris, tricine or citrate, 200 µg/ml bovine serum albumin, 2.5 mM $MgCl_2$, 200 µM dNTP (each) and either 16 mM $(NH_4)_2SO_4$ or 50 mM potassium acetate. Formamide at 10% concentration was used as indicated (see Enzyme/buffer notation). PCR buffers were made as 10× stocks requiring the addition of formamide as needed, dNTPs and the oligonucleotide primers and template DNA.

Enzyme/buffer notation

Test PCR buffers are named to indicate the presence of one or more components: Tris/potassium acetate, buffer A; Tris/ammonium sulfate, buffer B; tricine/ammonium sulfate, buffer D; EPPS/potassium sulfate, buffer E; EPPS/ammonium sulfate, buffer F; citrate/ammonium sulfate, buffer G. Component concentrations are described above.

Amplification of p53 exon 7 from genomic DNA

Part of p53 exon 7 surrounding codon 248 was amplified. The upstream primer (5′-GCCTCATCTTGGGCCTGTGTTATC-3′) hybridized within the preceding intron and the downstream primer (5′-GTGGATGGGTAGTAGTATGGAAGAAATC-3′) hybridized within exon 7. All PCR, RE digestion and ligation steps described throughout were performed using a GeneAmp PCR System 2400 (Perkin Elmer). Several buffers and enzymes were used (see PCR polymerases and buffers) as indicated. The p53 exon 7 amplification from genomic DNA was performed starting with a 20 µl reaction mixture containing 50 ng of DNA, 2.5 mM each dNTP and 12.5 pmol of each primer in 1× buffer without polymerase. The reaction mixture was covered with paraffin oil and preincubated for at least 1.5 min at 94°C in order to perform hot start by adding 1 µl of polymerase diluted in 1× buffer to introduce the required units of polymerase. The exon 7 segment was amplified for 40 cycles of 94°C for 15 s, 65°C for 2 min, with an additional 5 min at 65°C at the end of the last cycle. PCR amplifications departing from this procedure were performed as indicated.

PCR/RE/LDR

Fidelity assay. Templates were amplified with conversion primer pairs bracketing the central two base pairs of the *Msp*I site (CCGG) at codon 248 (Fig. 2B). Tubes were prepared containing 10 fmol/reaction of either PCR-amplified p53 exon 7 or wild-type synthetic duplex template, PCR buffer and primers. In parallel reactions, a synthetic 50 bp duplex marker template (MK), with the sequence CGGG replacing the *Msp*I site at codon 248, was added at 10^{-3}, 10^{-4}, 10^{-5} and 0 molar ratio to wild-type template. Reactions were preincubated for at least 1.5 min at 94°C with all components present in CiNF buffer except Vent(exo–) polymerase. A 'hot start' was performed by adding 1 µl of polymerase at 94°C. When preconversion was performed, two cycles of 94°C for 15 s, 55°C for 1 min, 60°C for 1 min were executed with 500 fmol each of the primers p53-248Q_6 and p53-248Q_6R. Afterwards, 1 pmol of p53Taq248T and p53Taq248TR primers were added. When preconversion was not performed, the reactions contained 1 pmol each of the primers p53Taq248T and p53Taq248TR or the control primers p53Msp248C and p53Msp248CR. After reactions with and without preconversion were performed, conversion PCR was carried out as follows: 5 cycles of 94°C for 15 s, 55 + 1°C/cycle for 1 min (temperature ramp), 60°C for 1 min; then 20 cycles of 94°C for 15 s, 60°C for 2 min; a final 60°C for 5 min extension. After three cycles of the temperature ramp 10 pmol of long zipcode conversion primers (p53zip248T and p53zip248TR or p53zip248C and p53zip248CR) were added. After conversion, the wild-type DNA was digested periodically during 20 cycles of 'zipcode' PCR (described below). Polymerase was inactivated by freezing and thawing twice. Finally, LDR was performed to detect the

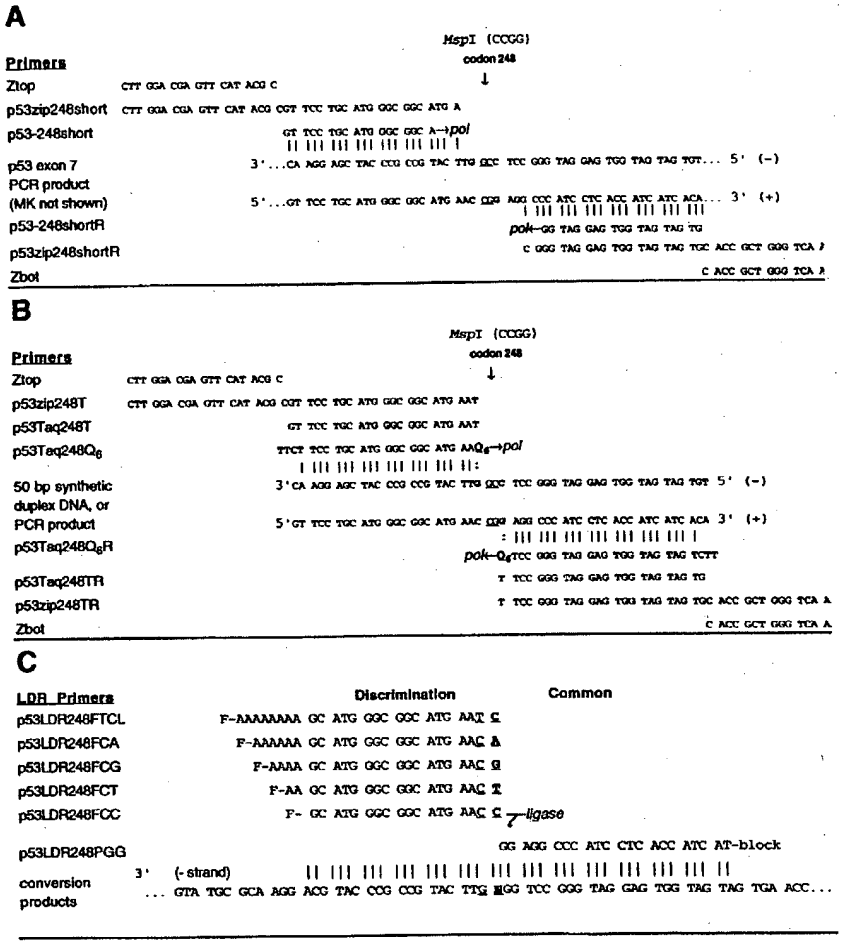


Figure 2. Primers used in PCR/RE/LDR. Complimentary (-- strand) sequences are shown in reverse orientation (3′→5′), in particular reverse strand primers (having names ending in R). **(A)** PCR fidelity assay. A synthetic 50 bp duplex marker template (MK) and wild-type p53 exon 7 PCR product are mixed at known ratios in parallel reactions. Perfect match primers p53-248short and p53-248shortR amplify the wild-type CCGG and marker CGGG. Then, longer zipcode-containing primers p53zip248short and p53zip248shortR were added. Finally, wild-type was repeatedly digested and reamplified with zipcode primers (Ztop and Zbot). **(B)** Preconversion was performed using primers containing 3′ convertide, e.g. p53-248Q_6. Conversion of the *Msp*I site to a *Taq*I site with or without preconversion was performed using 3′ natural base primers p53zip248T and p53zip248TR. Long primers were added as above and conversion products further amplified. Wild-type products were digested with the RE appropriate for the new site. Mutant products were preferentially amplified with zipcode primers. **(C)** LDR primer sets were designed to query the template sequence around the point of ligation. Perfectly hybridized upstream and downstream LDR primers with no overlap or gap are preferentially ligated with high specificity. Discrimination primers have different length 5′ tails to allow specific product separation on an acrylamide gel. Primers shown were used for identification of mutations occurring in the second base of the *Msp*I site (no conversion). An extra primer (p53LDR248FTCL) was used to compare C→T transitions at the first base and second base of the *Msp*I site. A comparable set of discrimination and common primers, used to identify mutations at the second base of the *Taq*I site in conversion products, had a T at the 3′ penultimate base in the discrimination primers and A at the 5′ penultimate base in the common primer.

conversion products without contribution from the original template (except in non-conversion control reactions).

'Zipcode' PCR. Wild-type sequences or wild-type conversion products were removed by restriction digestion. The appropriate RE was added to the reaction tube and supplemented with additional $MgCl_2$ as required to allow efficient digestion. *Msp*I digestion was performed at 37°C for 15 min using no additional $MgCl_2$, except when using citrate buffer. *Taq*I digestion was performed at 65°C for 30 min at 6 mM Mg^{2+} by adding 1 µl of enzyme diluted in 140 mM $MgCl_2$. The undigested conversion products were reamplified from 1 µl of a 10× dilution added to a 20 µl PCR reaction containing 10 pmol of the 'zipcode' primers Ztop and Zbot (Fig. 2B). These zipcode primers each contain a DNA sequence that is not similar in sequence to any genomic sequences present in the sample, thus only the products of previous PCR using primers containing the zipcode sequences will be efficiently amplified. Conversion products were amplified using Expand polymerase mix and buffer (see PCR polymerases and buffers). After an initial RE digest, zipcode PCR reamplification followed by redigestion was performed as follows: reactions were

preincubated at 94°C for at least 1.5 min then initiated with a hot start by adding 0.1 µl of RE-digested sample (1 µl of a 10× dilution) to a 20 µl reaction; 10 cycles of 94°C for 15 s, 65°C for 2 min. Zipcode PCR amplification products were redigested as described above.

Ligase detection reaction. Ligase detection reactions were performed in standard LDR buffer (25 mM Tris pH 7.6, 12 mM $MgCl_2$, 65 µg/ml bovine serum albumin, 100 mM KCl and 10 mM DTT). Each 20 µl reaction contained ~500 fmol of dsDNA (1 µl of PCR sample), 500 fmol of each discrimination primer and 750 fmol of common primer (Fig. 2C). Sets of discrimination and common primers were synthesized to detect the expected conversion products, i.e. converted to CNGG or TNGA at the *Msp*I position. The common primer was synthesized using 3′-Spacer C3 CPG columns and the 5′-end was phosphorylated on the column using phosphorylation reagent (see Oligonucleotide synthesis). Discrimination primers of each set varied at the 3′-terminal base to query the base in that location, i.e. the second base of the *Msp*I position. Discrimination primers had 5′ tails of different length and a FAM label for fluorescence detection. The tail size identified the primer and allowed physical separation of different LDR products on an acrylamide gel.

The LDR reaction was preincubated for 1.5 min at 94°C prior to the addition of 5 nmol *Tth* ligase enzyme under a layer of mineral oil. We used 10 LDR cycles of 94°C for 15 s, 65°C for 2 min. The reactions were then held at 94°C until cold quenched on ice and stored at –70°C. The LDR products were separated on 10% acrylamide gels containing 7 M urea with 0.6× TBE (1× TBE contains 90 mM Tris base, 90 mM borate, 2 mM EDTA) used in the gel and for the running buffer. Data were collected using an ABI 373 automated DNA sequencer and Applied Biosystems Genescan 672 software (GS Collection and GS Analysis).

Image processing

Raw gel pictures were produced by the ABI GS Analysis software. Dye-specific pictures were opened in Adobe Photoshop 3.0, cropped, resized and converted to grayscale. The grayscale images were opened in NIH Image 1.59, inverted and 1D vertical background was subtracted. Optionally, NIH Image could render a three-dimensional plot from a corrected two-dimensional picture. Background corrected pictures were reinverted and rendered in pseudocolor by Photoshop by replacing the color table to make subtle intensity differences easier to compare. Except for color replacement, only linear image processing was performed in order to preserve relative intensities.

RESULTS AND DISCUSSION

We developed PCR/RE/LDR to detect and identify low abundance mutations occurring within the *Msp*I site (CCGG) at codon 248 in the p53 gene (Fig. 1). An initial PCR amplifies exon 7 from genomic DNA. This product serves as the template for a second PCR that amplifies the central CpG dinucleotide in the *Msp*I site. To generate a restriction site in sequence lacking a pre-existing site, mismatch primers are used to alter one or more bases flanking the CpG dinucleotide. This results in a conversion PCR that creates a restriction site (NCGN→TCGA *Taq*I site, for example). In a generalized method for introducing contiguous Type II restriction sites, conversion PCR primers by necessity have 3′-terminal mismatches. To avoid unfavorable natural base mismatches that may result in insertion of an erroneous base at the next site (18,30), preconversion with 3′ nucleotide analog primers is performed. However, extension with 3′ analog primers produces a pool of degenerate products (19). Thus, after this preconversion step, natural base primers are used to selectively amplify the desired products.

We assessed mismatch conversion error relative to PCR error by performing parallel non-conversion control reactions and true conversion reactions with and without preconversion. Non-conversion reaction products retained the *Msp*I site (CCGG), while conversion introduces a *Taq*I site (TCGA). All PCR/RE/LDR steps were performed under similar conditions, varying only the primers and RE (*Msp*I or *Taq*I). In both cases, non-cleavable DNA is preferentially amplified. When wild-type DNA is selectively removed by digestion, it is necessary to determine the proportion of DNA with incorrect sequence produced relative to the initial quantity of DNA in the sample, which is nearly 100% wild-type. Parallel reactions were performed in which known fractions of MK DNA were present. The MK DNA contained a single base change in the *Msp*I site (C<u>G</u>GG), rendering it uncleavable by *Msp*I. C→G transversions are unlikely to occur through polymerase error. The MK standard curve allows quantification of mutations detected by LDR. PCR conditions were tested to minimize PCR error (observed in the non-conversion reactions) and mismatch extension errors (additional errors observed in the conversion reactions).

We tested various proofreading and non-proofreading polymerases, as different polymerase properties are required during target amplification from genomic DNA, conversion and reamplification steps in PCR/RE/LDR. Since it is essential throughout PCR/RE/LDR to minimize any alteration of the bases assayed by LDR, proofreading polymerases might seem the logical choice for maintaining the highest fidelity (31), however, they may interfere with conversion by mismatch primer extension. Hence, PCR conditions must be found which maximize the fidelity of non-proofreading polymerases (32).

Initially, we used PCR/RE/LDR as a high sensitivity assay to determine PCR conditions that maintain the highest fidelity throughout the procedure. Two main sources of error were expected: (i) polymerase misincorporation; (ii) DNA template degradation. Raising the PCR buffer pH improves long PCR, probably by decreasing depurination which leads to strand cleavage (22–24). While higher pH may decrease template damage, higher pH is also known to adversely affect polymerase fidelity (25–27). Therefore, we tested tricine, EPPS and citrate buffers which have pK_a values in the range 7–8 and $|\Delta pK_a|$ lower than Tris (see Materials and Methods). Tris cannot meet the dual constraints of neutral pH at high temperature to maintain template integrity and lower pH at the extension temperature to maintain polymerase fidelity, although most PCR fidelity and long PCR studies use Tris. Some investigators have explored the use of alternative buffers with lower $|\Delta pK_a|$ (25,26,28). Buffer-specific effects on PCR were tested by making PCR buffers containing identical components except for the buffering compound. We tested salt effects by making one set of test PCR buffers with ammonium sulfate and another with potassium acetate. The ΔpK_a of each buffer was determined in pure solution and in 1× PCR buffer mixtures (data not shown). Our results agreed with published ΔpK_a values of pure buffers (33,34) corrected by a small constant (0.005 pH units/°C), possibly due to a temperature

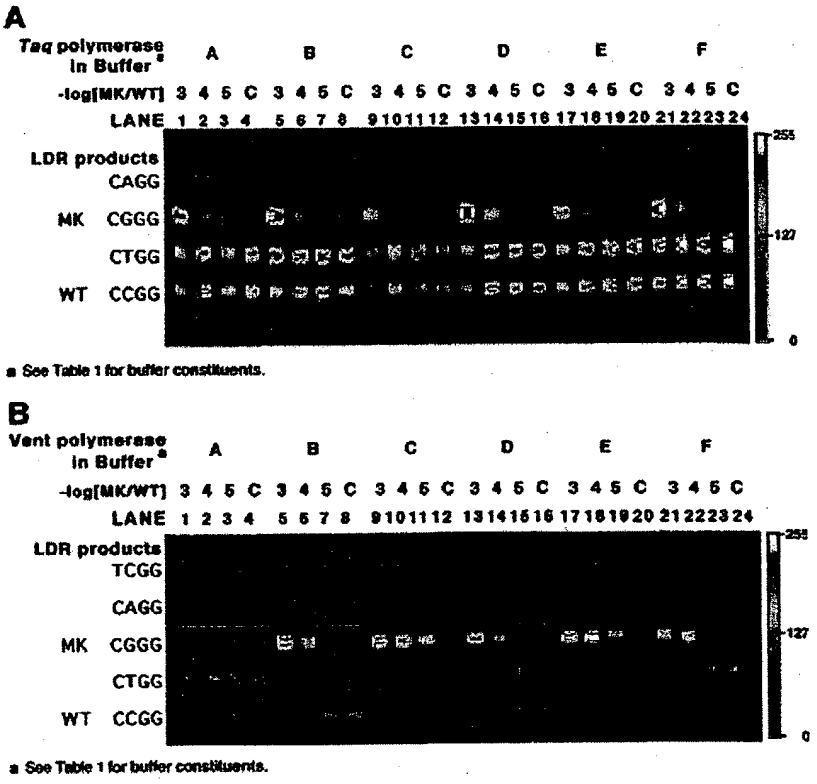


Figure 3. Buffer- and enzyme-dependent PCR errors detected by PCR/RE/LDR. The indicated polymerase/buffer combinations were used to amplify p53 exon 7 from genomic DNA. The same buffers were used in reactions with perfect match primers to reamplify the *Msp*I site. (A) *Taq* polymerase used in various test PCR buffers. (B) Vent polymerase used in various test buffers. Vent polymerase did not amplify p53 exon 7 from genomic DNA in TsK buffer. In this case only, two different enzyme/buffer sets were used for preamplification and 'conversion' (not actual conversion, since perfect match primers were used). The Ampli*Taq*/TsK exon 7 genomic DNA PCR product was substituted in the Vent/TsK reamplification. C indicates no MK was added (control reaction).

dependence of the pH probe itself. We adjusted the pH of our test PCR buffers to produce approximately neutral pH at 65°C. However, the 1× PCR buffers had somewhat different ΔpK_a values compared to the pure buffers; for example, 1× TsN had $\Delta pK_a = -0.033$/°C versus –0.030/°C for 100 mM Tris and 1× TcK had $\Delta pK_a = -0.022$/°C versus –0.025/°C for 100 mM tricine.

Test PCR buffers containing Tris, tricine or EPPS were used to test PCR fidelity with no conversion of the *Msp*I site (CCGG) at codon 248 of p53 (Fig. 3). Our objective in this experiment was to test the error rate of PCR using various buffers and polymerase enzymes. Since introduced errors create template that cannot be cleaved by the selected restriction enzyme, false positives accumulate as this error template continues to amplify alongside true mutant DNA. This experiment establishes the conditions necessary to achieve amplification while minimizing error. The same polymerase and buffer set was used in both preamplification of p53 exon 7 from genomic DNA and in the 'conversion' step. As mentioned, the 'conversion' step maintains the *Msp*I site by using perfectly matched primers whose 3′-ends terminate on the C and G bases flanking the central CpG. After an initial *Msp*I digest, template and amplification products were periodically redigested every 10 cycles during reamplification to remove WT sequence. Synthetic marker mutant MK with the sequence CGGG was present in these reactions at 10^{-3}, 10^{-4}, 10^{-5} or 0 ratio to wild-type (WT). MK will not be cleaved by *Msp*I restriction digestion, but will amplify with each PCR cycle to provide an internal control to measure product quantitities (see below). The MK product will also maintain its sequence, as the perfect match primers in the conversion step will again terminate on the C and G bases flanking the central GG. Error products resulting from MK PCR will in general lack *Msp*I sites and will be indistinguishable from regular MK template. If an *Msp*I site is accidentally created, the product will be destroyed by digestion.

For each buffer, LDR detected MK products in each of the four parallel reactions, with the 0 MK control indicating the background level of CGGG error produced. The intensities of other error products detected by LDR were compared to MK to estimate the fraction of each error product generated. Ampli*Taq* generated few transversions (C→G or C→A), but a large amount of C→T transition was observed (Fig. 3A). Vent generated much less of the C→T transition compared to Ampli*Taq* (Fig. 3B). Ampli*Taq* showed little dependence on the presence of potassium acetate in buffers A, C and E (Fig. 3A, lanes 1–4, 9–12 and 17–20) versus ammonium sulfate in buffers B, D and F (lanes 5–8, 13–16 and

21–24). Vent polymerase amplified template more efficiently in Tris/ammonium sulfate buffer B than Tris/potassium acetate buffer A (Fig. 3B, lanes 1–4 versus lanes 5–8), as described previously (22,31,35,36). However, Vent exhibited improved fidelity in tricine/potassium acetate buffer C (lanes 9–12) and EPPS buffer E (lanes 17–20) compared to tricine/ammonium sulfate buffer D (lanes 13–16) and EPPS/ammonium sulfate buffer F (lanes 21–24).

The relative fidelities of the different polymerase–buffer combinations may be described by their 'sensitivity' expressed as the $-\log_{10}$ of the ratio of MK to WT initially present. The C→T error for Ampli*Taq* amplification in Tris/potassium acetate buffer A can be taken as an example. If the signal for the CTGG error product (Fig. 3A, lane 2) is compared to the MK CGGG signal (Fig. 3A, lanes 1–3), the intensity of the signal most resembles the 10^{-3} MK:WT dilution (Fig. 3A, lane 1). Thus, the C→T error rate is 10^{-3}; the sensitivity is 3, since $-\log[MK/WT] = -\log[10^{-3}] = 3$. From this it can be seen that the higher the sensitivity, the lower the error rate. Reactions with higher sensitivities for each mutation had the highest overall fidelity (results summarized in Table 1). Many of the Vent reactions had sensitivities of 1 in 10^5 for every mutation (Fig. 3B), while the Ampli*Taq* reactions had sensitivities of 1 in 10^3 (Fig. 3A). Sensitivity indicates the usefulness of the assay rather than the error rate of the polymerase. Error (ER) per base per cycle may be estimated from the fraction (F) of all mutations occurring at one base which accumulated over 65 cycles (D) before digestion. For our purposes, the number of cycles is an estimate of the number of duplications, since multiple non-saturating PCRs were performed. From ER = F/D, Vent polymerase had an error rate of $<1 \times 10^{-7}$/base/cycle in tricine/potassium acetate buffer C, $\sim 2 \times 10^{-7}$/base/cycle in tricine/ammonium sulfate buffer D and 2×10^{-6}/base/cycle in TcN buffer. We observed an error rate of 2×10^{-5}/base/cycle mainly due to the C→T transition for Ampli*Taq* in Tris/potassium acetate buffer A. Elimination of this artifact could improve Ampli*Taq* fidelity by more than 10-fold. Others have used cloning and screening methods to estimate polymerase error (25,27,36,37) and denaturing gradient gel electrophoresis (DGGE) has also been used (31,32,35,38). However, these methods do not directly measure mutated DNA and do not detect all mutations. By cloning and DGGE methods, Vent polymerase has an error rate estimated as from 0.3 to 4×10^{-5}/base/cycle (27,35,38). The error rate of *Taq* polymerase has been estimated as from 0.8 to 9×10^{-5}/base/cycle (25,27,28), comparable to the error rate we observed for Ampli*Taq* in TsK buffer. Of the thermostable polymerases, *Pfu* has the lowest reported error rate estimated as from 0.7 to 1.6×10^{-6}/base/cycle (27,39,40). *Pfu* polymerase may also exhibit improved fidelity in tricine or other low $|\Delta pK_a|$ buffers.

While high fidelity proofreading enzymes appeared to improve amplification from genomic DNA, proofreading still must be avoided in the conversion step. We tested different high fidelity genomic amplification conditions in combination with fixed conversion conditions. Genomic amplification was performed with either Vent(exo–) in citrate/ammonium sulfate buffer G or Vent(exo–) in citrate/ammonium sulfate buffer G with 10% formamide (Table 1). Non-conversion primers were used with Vent(exo–) to optimize PCR fidelity in anticipation of conversion by mismatch primer extension. Our highest fidelity conditions were as follows. Genomic amplifications with Vent/buffer G were initiated by spiking genomic amplification product from Expand/buffer C with 10% formamide reactions after three cycles. These Vent/buffer G reactions required 4 mM Mg^{2+} and PCR primers, but no additional genomic DNA was provided (see Table 1 for observed error rate with other conditions tested).

Table 1. Comparison of fidelity using proofreading and non-proofreading polymerases in different buffers for PCR to amplify the target sequence from genomic DNA and for conversion PCR

Polymerase enzymes (1) Genomic DNA PCR→(2) conversion	Buffer (1)→(2)	Limiting error	Error rate Total	Per cycle[a]
Taq→*Taq*	A→A	C→T	10^{-3}	2×10^{-5}
Taq→*Taq*	B→B	C→T	10^{-3}	2×10^{-5}
Taq→*Taq*	C→C	C→T	10^{-3}	2×10^{-5}
Taq→*Taq*	D→D	C→T	10^{-3}	2×10^{-5}
Taq→*Taq*	E→E	C→T	10^{-3}	2×10^{-5}
Taq→*Taq*	F→F	C→T	10^{-3}	2×10^{-5}
Taq[b]→Vent	A→A	C→T	$>10^{-3}$	$>2 \times 10^{-5}$
Vent→Vent	B→B	C→T	10^{-5}	2×10^{-7}
Vent→Vent	C→C	C→T	$<10^{-5}$	$<2 \times 10^{-7}$
Vent→Vent	D→D	C→T	10^{-4}	2×10^{-6}
Vent→Vent	E→E	C→T	$<10^{-5}$	$<2 \times 10^{-7}$
Vent→Vent	F→F	C→T	10^{-5}	2×10^{-7}
Vent(exo–)→Vent	C→G	C→T	10^{-4}	2×10^{-6}
Vent(exo–)→Vent(exo–)	C→G	C→T	10^{-3}	2×10^{-5}
Vent→Vent(exo–)	C→G	C→T	10^{-4}	2×10^{-6}
Vent→Vent(exo–)	C→G(f)	C→T	10^{-5}	2×10^{-7}
Vent[c]→Vent(exo–)	G (4)→G	C→T	10^{-4}	2×10^{-6}
Vent→Vent(exo–)	G (8)→G(f)	C→T	10^{-5}	2×10^{-7}
Vent[c]→Vent(exo–)	G (4)→G(f)	C→T	$<10^{-5}$	$<2 \times 10^{-7}$

[a]Based on a minimum of 50 total cycles, i.e. observed error/50.

[b]No Vent PCR product from genomic DNA. *Taq* amplified product used for Vent conversion PCR.

[c]Template added by taking 1 µl after the third PCR cycle from a parallel genomic DNA amplification using Expand polymerase mix in buffer C.

Taq and Vent polymerases were initially tested using one buffer for genomic amplification and conversion. During the conversion step, only non-conversion of the *Msp*I site near p53 codon 248 was performed using short perfect match primers (Fig. 2A) to determine the background level of polymerase error. LDR quantified *Msp*I site mutations at the second position (CCGG→CNGG). Fidelity was compared in parallel reactions using proofreading and non-proofreading polymerases in genomic amplification and conversion. Expand polymerase mix, *Taq* with proofreading *Pfu* polymerase added, was used to initiate target amplification from genomic DNA for subsequent high fidelity Vent polymerase PCR. Vent polymerase was substituted with non-proofreading Vent(exo–) to determine whether proofreading was required and also in the conversion step where proofreading is not permitted. The effect of 10% formamide in the conversion PCR buffer was also tested. All buffers contained 200 µg/ml bovine serum albumin, 2.5 mM $MgCl_2$ and 200 µM dNTP (each). Specific components were: A (TsK), 20 mM Tris pH 9.1, 50 mM potassium acetate (standard *Taq* polymerase buffer); B (TsN), 20 mM Tris pH 9.1, 16 mM ammonium sulfate (standard Vent polymerase buffer); C (TcK), 20 mM tricine pH 8.7, 50 mM potassium acetate; D (TcN), 20 mM tricine pH 8.7, 16 mM ammonium sulfate; E (EpK), 20 mM EPPS pH 8.4, 50 mM potassium acetate; F (EpN), 20 mM EPPS pH 8.4, 16 mM ammonium sulfate; G (CiN), 20 mM citrate pH 7.6, 16 mM ammonium sulfate. (f), presence of 10% formamide; (4), increase to 4 mM $MgCl_2$; (8); increase to 8 mM $MgCl_2$.

We found PCR conditions for each step in PCR/RE/LDR that maintain high fidelity when no mismatch conversion was performed. With known high fidelity PCR conditions, we next tested conversion by changing the p53 codon 248 *Msp*I site

(CCGG) into a *Taq*I site (TCGA). MK (CGGG) was added as before in parallel reactions to measure fidelity relative to the initial wild-type DNA present. High fidelity PCR was performed as described above and some (but not all) reactions were subjected to preconversion. Preconversion was performed using primers containing the degenerate pyrimidine nucleotide analog Q_6 at their 3′-ends (Fig. 2B). The final conversion was accomplished using natural base 3′ T mismatch primers. Products were detected using LDR to interrogate the second base position in the *Msp*I, *Taq*I and MK sequence: CNGG or TNGA. Fidelity for conversion with and without preconversion was compared to a non-conversion control. Successful conversion will change the *Msp*I site (CCGG) into a *Taq*I site (TCGA); MK will also be converted from CGGG to TGGA. However, the main issue of conversion success is the maintenance of the central bases in all cases: CpG for *Taq*I conversions and GpG for MK. Figure 4 shows the results of conversion. Lanes 1–4 (C:G) are non-converted reactions that were digested with *Msp*I; lanes 5–8 (Q_6:G) are preconverted/converted reactions that were digested with *Taq*I; lanes 9–12 (T:G) are converted reactions lacking preconversion that were digested with *Taq*I. PCR/RE/LDR with no conversion was sensitive to better than 1 in 10^4 using the previously determined best conditions for preamplification and conversion (Fig. 4, C:G, lanes 1–4). PCR/RE/LDR with conversion of the *Msp*I site to a *Taq*I site by T mismatch primers was apparently very successful at first glance (T:G, lanes 9–12). As would be expected for successful conversion, no *Msp*I product can be detected in the _CG_ region of the figure, hence, it appears that the site was converted to *Taq*I and then digested. However, although a very large fraction of MK (CGGG) is observed in the reactions with added MK (T:G, lanes 9–11), the same large fraction is also observed in the 0 MK control lane (T:G, lane 12). Thus, the entire quantity of MK is an artifact produced by mismatch extension of the natural base T primers. This event would convert the second position C in the *Msp*I site to a G during extension, mimicking the internal sequence of MK (CCGG→TGGA). Preconversion with Q_6 primers eliminates the MK artifact (Q_6:G, lanes 5–8).

The greater amount of WT present in non-converted samples (Fig. 4, lanes 1–4) compared to Q_6 converted samples (lanes 5–8) may be due to inhibition of *Msp*I digestion by formamide. Formamide apparently inhibits *Msp*I digestion as evidenced by the presence of strong wild-type LDR bands (WT) in the non-conversion control (C:G lanes), which are not present after *Taq*I digestion of the converted sequence (Q_6:G and T:G lanes).

The low amount of MK product seen in the Q_6 10^{-4} and 10^{-5} MK lanes (Fig. 4, Q_6:G, lanes 6 and 7) compared to the respective non-conversion control reactions (C:G, lanes 2 and 3) may be due to low efficiency of MK conversion. The production of a *Taq*I site actually requires two conversions, one on each side of the central CpG dinucleotide. Lowering the concentration of MK 10-fold may reduce MK conversion far more than 10-fold. Thus, with only one side of the MK sequence converted in a large amount of its product, one half of the LDR primers will be unable to properly hybridize to this sequence and ligation will not occur. LDR detection will only reveal the lesser quantity of fully converted template. Nevertheless, the amount of MK product is greater than the control in these two lanes (compare Fig. 4 lane 8 to lanes 6 and 7). While formamide may reduce conversion efficiency, conversion fidelity is greatly improved.

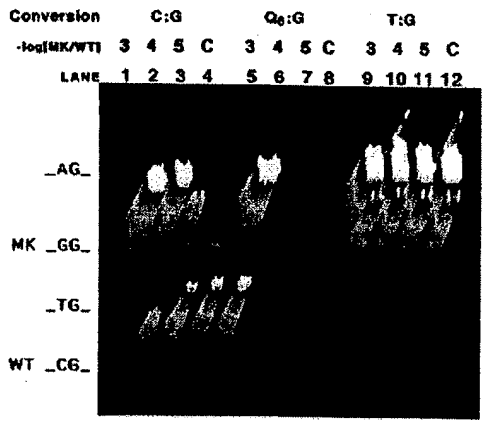


Figure 4. Comparison of conversion fidelity. The relative intensities of conversion reaction products is indicated by color and the height of each peak in a 3-dimensional plot (Materials and Methods). Marker (MK) DNA (with CGGG replacing the *Msp*I site) was added at known ratios to wild-type (WT) in parallel reactions. The –log(MK:WT) indicates relative fraction of MK present, e.g. –log(MK:WT) = 3 means the ratio of MK to WT was 1:1000. C indicates no MK was added (control reaction). Non-conversion control reactions (C:G) were performed using perfect match 3′ C primers. Conversion of the *Msp*I site (CCGG) to a *Taq*I site (TCGA) was performed using natural base 3′ T primers with and without preconversion using 3′ Q_6 nucleotide analog primers (Q_6:G and T:G reactions, respectively). LDR products from *Msp*I non-conversion contain CNGG and products from *Taq*I conversion contain TNGA, but only the the central bases (second and third bases) are indicated as _NG_. The LDR products were designed to separate on acrylamide gels by two base differences in size. Some undetermined bands of intermediate size were also observed. Lanes 1–4 were digested with *Msp*I, while lanes 5–12 were digested with *Taq*I during PCR/RE/LDR.

Newton *et al.* found that C·T, A·A and T·T mismatches were all far more difficult to extend with *Taq* polymerase than purine·pyrimidine mismatches (7). These results reflect PCR efficiency of extension rather than fidelity. Others have observed low fidelity in extending natural base mismatches (18,30). Use of a nucleotide analog with structural similarities to multiple bases could potentially be used to allow polymerase extension (reading) from the analog when paired with different bases and insertion of different bases opposite the analog (writing). For the purposes of this assay, the efficiency of the process need not necessarily be high. However, successful conversion requires high PCR fidelity to ensure that only the bases targeted for conversion are altered. False positive mutation artifacts will result from alterations of bases not targeted for conversion within the sequence probed for mutations. Preconversion using 3′ Q_6 primers forming a Q_6:G mismatch avoids starting polymerase extension with a G·T mismatch. In subsequent amplification cycles, A is apparently written frequently opposite Q_6. This observation is consistent with the results of Hill *et al.* in which Q_6 base paired like C and T with nearly equal frequency (41). Facile tautomerization allows Q_6 to mimic either pyrimidine when base paired and avoids mismatch wobble. When the natural base primer is added after preconversion, a significant quantity of perfect match template already exists, otherwise the MK artifact would appear in the reaction regardless of preconversion. Other nucleotide analogs in addition to Q_6 may serve as a bridge for more efficient conversions (19).

We have measured the fidelity of polymerase extension from primers in PCR and found conditions which in some cases improve fidelity. Presumably, higher fidelity resulted from a decrease in polymerase misincorporation, primer slippage and template degradation. PCR/RE/LDR allows the measurement of very low level 'mutant' sequences by preferentially amplifying non-wild-type sequences. Our results clearly demonstrate that natural base mismatch primer extension cannot be used as a general technique to create restriction sites at will in any sequence for RFLP analysis. As demonstrated here (Fig. 4) and observed previously (18,30), natural base mismatch extension is prone to error. To perfectly engineer a restriction site from existing sequence, an error-free approach is required. Our results indicate that the use of nucleotide analogs combined with high fidelity PCR conditions may radically decrease errors. Monitoring the true specificity of primer extension was possible in these studies because LDR can measure specific PCR errors accurately and with high sensitivity. Thus, the products of different polymerases and buffers could be assayed at different steps during PCR/RE/LDR to maximize both PCR efficiency and fidelity. As a result, a PCR/RE/LDR strategy could be assembled to achieve the goal of 10^5 sensitivity. However, this highest sensitivity was achieved only in the special case of no conversion at a pre-existing *MspI* site. At this time, primer slippage remains an important mechanism through which mismatch primer extension errors can arise (19). Although the importance of this source of error *in vivo* is uncertain, it may have a dramatic impact on allele-specific PCR and other *in vitro* methods of mutation detection. An additional source of error arises from using natural base primers to select specific sequences for amplification following preconversion with nucleotide analogs. This is because a fraction of the selective natural base primers may form a mismatched pair with bases other than the intended base. It is known that a characteristic set of different bases insert opposite nucleotide analogs (19,41). Thus, a high fidelity mismatch primer extension protocol awaits the development of new convertides that can overcome these problems. In combination with high fidelity PCR and LDR monitoring of efficiency, mismatch primer extension may become a technique for the precise introduction of desired mutations without artifacts.

ACKNOWLEDGEMENTS

We thank Michael Wigler, Thierry Soussi, Mark Sobel, Jerard Hurwitz, Saul Silverstein, Harry Ostrer, Michael Osborne, Daniel Knowles, Pat Paty, Reyna Favis and members of the Barany and Paty laboratories for technical assistance and helpful discussion. This work was supported by grants from the National Cancer Institute (CA65930-02), Strang Cancer Prevention Center and the Applied Biosystems Division of the Perkin Elmer Corporation.

REFERENCES

1 Cotton,R.G. (1993) *Mutat. Res.*, **285**, 125–144.
2 Bottema,C.D. and Sommer,S.S. (1993) *Mutat. Res.*, **288**, 93–102.
3 Fan,E., Levin,D.B., Glickman,B.W. and Logan,D.M. (1993) *Mutat. Res.*, **288**, 85–92.
4 Cariello,N.F. and Skopek,T.R. (1993) *Mutat. Res.*, **288**, 103–112.
5 Sidransky,D., Tokino,T., Hamilton,S.R., Kinzler,K.W., Levin,B., Frost,P. and Vogelstein,B. (1992) *Science*, **256**, 102–105.
6 Brennan,J.A., Mao,L., Hruban,R.H., Boyle,J.O., Eby,Y.J., Koch,W.M., Goodman,S.N. and Sidransky,D. (1995) *N. Engl. J. Med.*, **332**, 429–435.
7 Newton,C.R., Graham,A., Heptinstall,L.E., Powell,S.J., Summers,C., Kalsheker,N., Smith,J.C. and Markham,A.F. (1989) *Nucleic Acids Res.*, **17**, 2503–2516.
8 Tada,M., Omata,M., Kawai,S., Saisho,H., Ohto,M., Saiki,R.K. and Sninsky,J.J. (1993) *Cancer Res.*, **53**, 2472–2474.
9 de Fromentel,C.C. and Soussi,T. (1992) *Genes Chromosomes Cancer*, **4**, 1–15.
10 Barany,F. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 189–193.
11 Barany,F. (1991) *PCR Methods Applicat.*, **1**, 5–16.
12 Day,D.J., Speiser,P.W., White,P.C. and Barany,F. (1995) *Genomics*, **29**, 152–162.
13 Khanna,M., Park,P., Zirvi,M., Cao,W., Picon,A., Day,J., Paty,P. and Barany,F. (1999) *Oncogene*, **18**, 27–38.
14 Luo,J., Bergstrom,D.E. and Barany,F. (1996) *Nucleic Acids Res.*, **24**, 3071–3078.
15 Sandy,M.S., Chiocca,S.M. and Cerutti,P.A. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 890–894.
16 Pourzand,C. and Cerutti,P. (1993) *Mutat. Res.*, **288**, 113–121.
17 Kwok,S., Kellogg,D.E., McKinney,N., Spasic,D., Goda,L., Levenson,C. and Sninsky,J.J. (1990) *Nucleic Acids Res.*, **18**, 999–1005.
18 O'Dell,S.D., Humphries,S.E. and Day,I.N. (1996) *Genome Res.*, **6**, 558–568.
19 Day,J.P., Hammer,R.P., Bergstrom,D. and Barany,F. (1999) *Nucleic Acids Res.*, **8**, 1810–1818.
20 Kumar,R. and Barbacid,M. (1988) *Oncogene*, **3**, 647–651.
21 Anderson,J.A., Irish,J.C. and Ngan,B.Y. (1992) *J. Otolaryngol.*, **21**, 321–326.
22 Barnes,W.M. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 2216–2220.
23 Cheng,S., Fockler,C., Barnes,W.M. and Higuchi,R. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 5695–5699.
24 Sang,N., Condorelli,G., De Luca,A., MacLachlan,T.K. and Giordano,A. (1996) *Anal. Biochem.*, **233**, 142–144.
25 Eckert,K.A. and Kunkel,T.A. (1990) *Nucleic Acids Res.*, **18**, 3739–3744.
26 Eckert,K.A. and Kunkel,T.A. (1991) *PCR Methods Applicat.*, **1**, 17–24.
27 Cline,J., Braman,J.C. and Hogrefe,H.H. (1996) *Nucleic Acids Res.*, **24**, 3546–3551.
28 Brail,L., Fan,E., Levin,D.B. and Logan,D.M. (1993) *Mutat. Res.*, **303**, 171–175.
29 Cha,R.S., Zarbl,H., Keohavong,P. and Thilly,W.G. (1992) *PCR Methods Applicat.*, **2**, 14–20.
30 Eiken,H.G., Odland,E., Boman,H., Skjelkvale,L., Engebretsen,L.F. and Apold,J. (1991) *Nucleic Acids Res.*, **19**, 1427–1430.
31 Keohavong,P., Ling,L., Dias,C. and Thilly,W.G. (1993) *PCR Methods Applicat.*, **2**, 288–292.
32 Keohavong,P. and Thilly,W.G. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 9253–9257.
33 Good,N. (1976) In Fasman,G.D. (ed.) *Handbook of Biochemistry and Molecular Biology*. CRC Press, Cleveland, OH, pp. 367–369.
34 Blanchard,J.S. (1984) *Methods Enzymol.*, **104**, 404–414.
35 Cariello,N.F., Swenberg,J.A. and Skopek,T.R. (1991) *Nucleic Acids Res.*, **19**, 4193–4198.
36 Mattila,P., Korpela,J., Tenkanen,T. and Pitkanen,K. (1991) *Nucleic Acids Res.*, **19**, 4967–4973.
37 Huang,H. and Keohavong,P. (1996) *DNA Cell Biol.*, **15**, 589–594.
38 Ling,L.L., Keohavong,P., Dias,C. and Thilly,W.G. (1991) *PCR Methods Applicat.*, **1**, 63–69.
39 Lundberg,K.S., Shoemaker,D.D., Adams,M.W., Short,J.M., Sorge,J.A. and Mathur,E.J. (1991) *Gene*, **108**, 1–6.
40 Andre,P., Kim,A., Khrapko,K. and Thilly,W.G. (1997) *Genome Res.*, **7**, 843–852.
41 Hill,F., Loakes,D. and Brown,D.M. (1998) *Proc. Natl Acad. Sci. USA*, **95**, 4258–4263.

EXHIBIT 6

[CANCER RESEARCH 55, 1444–1447, April 1, 1995]

Advances in Brief

Increased Prevalence of K-*ras* Oncogene Mutations in Lung Adenocarcinoma[1]

Nancy E. Mills, Charles L. Fishman, William N. Rom, Neil Dubin, and Daniel R. Jacobson[2]

Department of Medicine, Divisions of Hematology [N. E. M., D. R. J.], Oncology [N. E. M.], Pulmonary and Critical Care Medicine [C. L. F., W. N. R.], and Environmental Medicine [N. D.] and Kaplan Comprehensive Cancer Center [D. R. J.], New York University Medical Center, New York, New York 10016

Abstract

Reported estimates of *ras* mutation prevalence in lung adenocarcinoma of 15–24% may be underestimates because of the insensitivity of the assays used. We have devised a rapid, non-radioactive assay for *ras* mutations, which detects 1 mutant allele/10^5 normal alleles and have used it to study DNA isolated from 53 lung tumor samples (including 28 adenocarcinomas) previously analyzed by PCR/allele specific oligonucleotide hybridization, which is less sensitive. We detected mutations in 13 of 28 samples, including 7 not detected by PCR/allele specific oligonucleotide hybridization. We also found *ras* mutations in 14 of 25 previously unstudied samples (56%). Our results indicate that the prevalence of K-*ras* codon 12 mutations in lung adenocarcinoma is higher than previously reported; thus, *ras* mutations may be more clinically useful as molecular markers for lung cancer than has been appreciated.

Introduction

Molecular tumor markers, including mutant *ras* alleles, may offer clinically useful tools for diagnostic and prognostic purposes in lung cancer. Several studies have reported that 15–24% of human lung adenocarcinomas contain mutations in K-*ras* codon 12 (1–4) and that *ras* mutations portend a poor prognosis in lung adenocarcinoma (2, 5, 6). A potential difficulty with these studies, however, lies in the insensitivity of the assays used to detect *ras* mutations. In the largest series of *ras* mutations in lung cancer (1), as well as in most other recent studies of *ras* mutations in human tumors, *ras* mutations have been detected by PCR/ASO-H,[3] which detects only mutations which are present in at least 10% of the total number of copies of the corresponding *ras* gene in the DNA sample (2, 7, 8). In tumors in which a small fraction of the malignant cells contain a *ras* mutation or in studies of DNA isolated from clinical samples containing malignant cells mixed with genetically normal cells, this limited assay sensitivity risks a high rate of false negative results.

To address this concern, we have used a simple, highly sensitive assay for K-*ras* codon 12 mutations to reexamine DNA samples isolated from lung adenocarcinomas which were previously reported as being *ras* mutation negative by PCR/ASO-H, as well as to study 25 previously unstudied samples. The detection of significant numbers of mutations which had been undetectable by PCR/ASO-H would indicate that *ras* mutations are of greater potential clinical utility as tumor markers than suggested by the reported prevalence of 15–24% and might also have an impact on prognostic studies. In addition, demonstration that many lung adenocarcinomas have *ras* mutations in only a small fraction of tumor cells would have implications for our understanding of the role of *ras* mutations in the process of lung carcinogenesis.

Materials and Methods

DNA samples isolated from 53 resected lung tumors (28 adenocarcinomas, 12 squamous cell carcinomas, 9 large cell carcinomas, 2 bronchoalveolar carcinomas, and 2 carcinoid tumors), previously analyzed for *ras* mutations by PCR/ASO-H by Rodenhuis and Slebos (1) at the Netherlands Cancer Institute, were reanalyzed for *ras* mutation status. The assay used, PCR-PIREMA, was adapted from a method which detected as little as 1 mutant allele in N-*ras* codon 12 or 61 per 10^6 normal alleles (9). PCR-PIREMA was performed blinded with regard to tumor histology and *ras* mutation status according to PCR/ASO-H. DNA isolated (10) from lung adenocarcinoma tissues obtained at the New York University Medical Center was phenol/chloroform extracted and desalted and concentrated to 5–10 μl in Microcon 100 concentrators (Amicon, Inc., Beverly, MA) twice, prior to PCR.

Mutations in K-*ras* codon 12 were detected as shown in Fig. 1. The most important modification from the method as described previously (9) was the use of greatly decreased concentrations of nucleotides and $MgCl_2$ in the first and second PCR reactions and lengthened annealing and synthesis times. A gel purification step (9) was also eliminated, which decreases assay sensitivity but simplifies the procedure. The modified protocol detects mutant alleles present at the level of 0.1%, as determined by titration of samples known to contain *ras* mutations (9). All samples were subjected to the entire PCR-PIREMA process at least twice. Extensive measures were taken to prevent cross-contamination of samples (9). Multiple normal controls and negative controls (no DNA in the PCR reaction) were included in all experiments.

Results

Among the 53 samples tested by PCR-PIREMA were 22 from adenocarcinomas which had been mutation negative by PCR/ASO-H (1). Of these 22, 7 (32%) were mutation positive by PCR-PIREMA (Fig. 2). The 6 adenocarcinomas which had been mutation positive by PCR/ASO-H were all positive for the same mutation by PCR-PIREMA (Table 1). Although DNA from the majority of samples is no longer available for study, by extrapolating from our data to the entire original group of 181 adenocarcinomas, we estimate that the true percentage which would be expected to be mutation-positive would be 24% (the percentage positive by PCR/ASO-H) (1) + 76% (the percentage negative by PCR/ASO-H) × 0.32, or a total of 48%. Of the 23 non-adenocarcinomas, all of which had been mutation negative by PCR/ASO-H, only one was found by PCR-PIREMA to contain a mutation (GTT in one large cell sample).

The K-*ras* codon 12 mutation status was also determined by PCR-PIREMA for 25 lung adenocarcinoma tissues from the New York University Medical Center. Fourteen samples (56%; 95% confidence interval, 35–76%) contained mutations (5 TGT, 4 GTT, 4 GAT, 1 AGT). The percentage of mutation-positive samples is not statistically different from the percentage in the Dutch group if a true positive percentage of 48% is assumed (2-tailed t test, $P = 0.55$) but is significantly greater than the 24% originally reported for the Dutch group ($P = 0.0026$).

Received 1/9/95; accepted 2/20/95.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] This work was supported by National Cancer Institute Grant NIH (NCI) K12 CA01713-02 (N. E. M.), American Cancer Society Institutional Grant IRG-14-35 (D. R. J.) and National Institute of Environmental Health Sciences Grant NIH (NIEHS) T32 ES07267-02 (C. L. F., W. N. R.).

[2] To whom requests for reprints should be addressed, at Research Service 151, New York V.A. Hospital, 423 E. 23 St., New York, NY 10010.

[3] The abbreviations used are: ASO-H, allele specific oligonucleotide hybridization; PCR-PIREMA, PCR-primer introduced restriction with enrichment for mutant alleles.

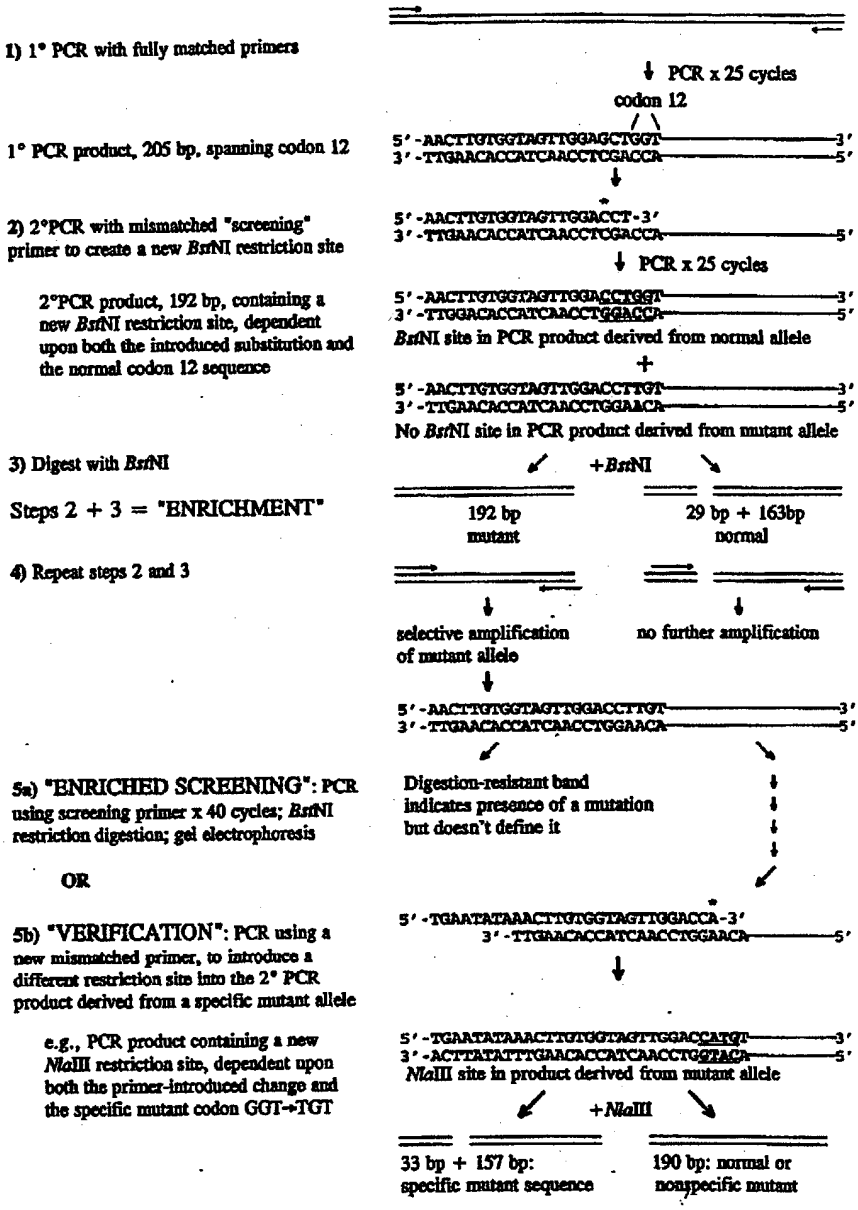


Fig. 1. PCR-PIREMA protocol. First PCR reactions contained 10 μM concentrations of each nucleotide, 1.2 mM $MgCl_2$, and fully matched primers flanking K-*ras* exon 1 (*Step 1*). Aliquots of the first PCR products were used in second PCR reactions containing 4 μM concentrations of each nucleotide and 1.5 mM $MgCl_2$, using a 5'-mismatched primer, to introduce a new *Bst*NI restriction site into PCR products derived from alleles containing the normal sequence at positions 1 and 2 of codon 12 (*Step 2*). The first and second PCR reactions were cycled at 94°C for 1 min, 55°C (for the first PCR) or 40°C (for the second PCR) for 2 min, and 74°C for 3 min. PCR products from *Step 2* were digested with *Bst*NI (*Step 3*). *Steps 2* and *3* were then repeated (*Step 4*) except on samples in which mutations which were strongly visible after a single round of enrichment. The digests were then amplified under "standard" PCR conditions (200 μM each nucleotide, 1.5 mM $MgCl_2$, 55°C annealing) using the same primers as in *Step 2* followed by repeat *Bst*NI digestion (*Step 5a*); these products were electrophoresed on 2.5% agarose gels and stained with ethidium bromide, with a digestion-resistant 192-base pair (*bp*) band indicating the presence of a K-*ras* codon 12 mutation. To identify the specific base substitution present, *Step 4* PCR products were PCR amplified with different mismatched 5'-primers which introduced new restriction sites into the PCR product, dependent upon both the specific mutations and the introduced substitution(s) (*Step 5b*, verification). PCR conditions for verification were 200 μM each nucleotide, 1.5 mM $MgCl_2$, 94°C for 30 s, 50°C for 1 min, 74°C for 1 min for 40 cycles. These PCR products were digested with the appropriate restriction enzymes and electrophoresed as above; the presence of a digested PCR product indicates the presence of a specific mutation. The PCR primers and restriction endonucleases used are as described previously (9), except that since the screening step now uses *Bst*NI rather than *Bsa*JI (*Step 2*), the fourth last base of all 5'-primers has been changed from C to A to eliminate an unnecessary mismatch. °, mismatched base.

Discussion

These data indicate that an assay which detects mutations present in 0.1% of alleles can detect K-*ras* codon 12 mutations in about 50% of lung adenocarcinomas. Our data demonstrate that the largest previously published report on *ras* mutations in lung adenocarcinoma (1) contained significant numbers of false negatives and thereby underestimated the prevalence of *ras* mutations in lung adenocarcinoma. Because other smaller series using similarly insensitive assays have reported mutation frequencies even less than 24%, they may have contained even higher percentages of false negatives.

Two factors, in conjunction with the use of an insensitive assay, could have contributed to difficulty in detecting mutations in these studies. First, in some cases, only a small fraction of tumor cells may contain the mutation. Since cancer is a clonal disease, one might expect that in mutation-positive cases, the mutation would be contained in all cells of the tumor. In fact, this has not to our knowledge been previously studied in lung adenocarcinoma; furthermore, in the one malignancy of which we are aware where this question has been examined (acute myelogenous leukemia), *ras* mutations are generally present in only a fraction of the malignant cells, and this fraction varies widely from case to case (9, 11–14). Second, clinical tissue samples obtained during surgical resection of lung masses generally contain both tumor cells and genetically normal cells; DNA isolated from such specimens will be derived from both populations. Thus, the samples from the Netherlands Cancer Institute included samples containing as few as 25% malignant cells (15). Similarly, one smaller study included some samples containing as few as 20% malignant cells (2); the latter group acknowledged the consequent likelihood of some false negatives. Both the presence of *ras* mutations in a minority of the malignant cells and the mixture of normal with malignant cells

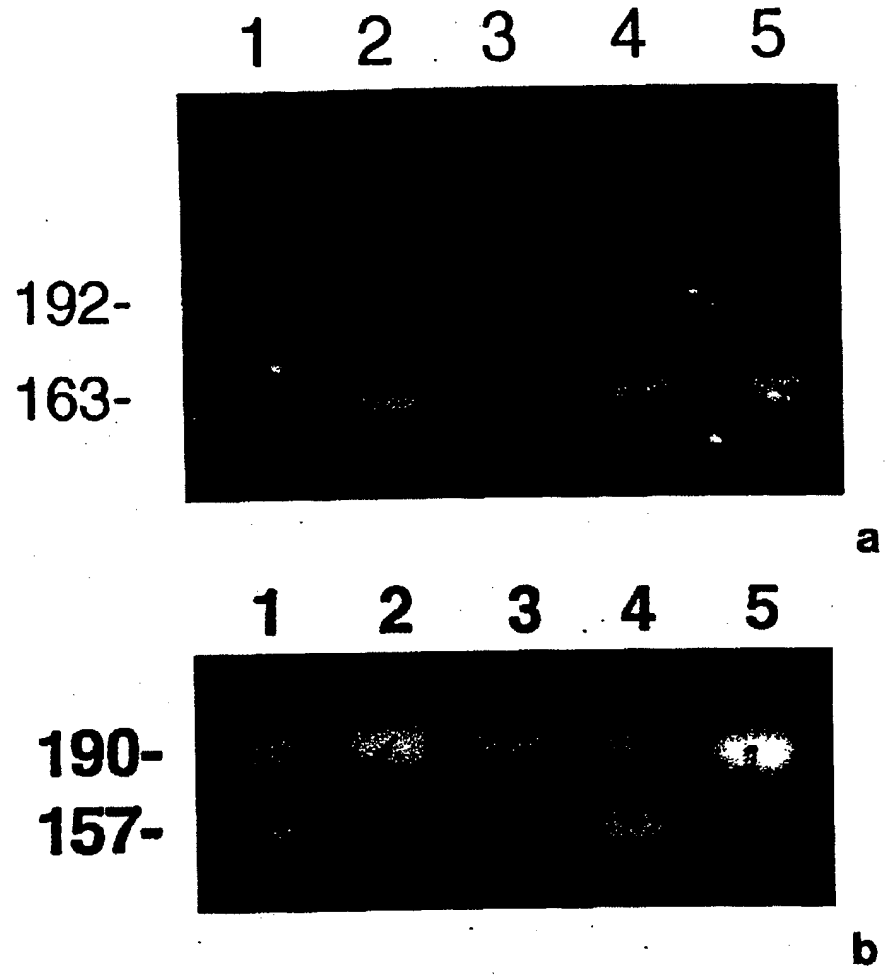


Fig. 2. PCR-PIREMA analysis of DNA samples from tumors previously analyzed by ASO-H. *a*, enriched screening (*Bst*NI digestion). *Lane 1*, undigested PCR product of 192 base pairs. *Lane 2*, normal control; *Lanes 3–5*, DNA samples from lung adenocarcinomas. The 192-base pair band in *Lanes 4* and *5* indicates the presence of a mutation in those samples. *b*, verification for TGT mutation by *Nla*III digestion. *Lanes 1–4*, adenocarcinoma samples; *Lane 5*, normal control. The digested 157-base pair band in *Lanes 1*, *3*, and *4* indicates the presence of a TGT mutation in those samples. For the 7 samples from adenocarcinomas that are positive for *ras* mutations by PCR-PIREMA but were mutation negative by ASO-H, the mutant genotypes are: GAT (4 samples); TGT (2 samples); and GTT (1 sample). These are the same mutations which predominated in the original analysis (1).

in clinical samples could lead to purified DNA samples in which the percentage of mutant alleles is less than the detection threshold of the assay used, *i.e.*, 10%.

Many genetic changes have been identified in lung cancer (16), but little is known about the chronology of their development. Some genetic changes may represent early activation events, while others are likely to occur later, as progression events related to invasion and metastasis. Limited evidence has been interpreted as suggesting that in some human tumor types, including lung cancer, *ras* mutations may fall into the former category (17). On the other hand, if significant numbers of cases of lung adenocarcinoma have *ras* mutations in only a small percentage of the cancer cells (as implied by the detection of these mutations by PCR-PIREMA but not by PCR/ASO-H), then this would suggest that the *ras* mutation was not an early, initiating event, but was more likely to represent a secondary event in a subclone of the tumor, as has been argued for acute myelogenous leukemia (18). This concept is consistent with the association of *ras* mutations with increased tumor growth and invasiveness, as suggested by the poorer prognosis of mutation-positive than mutation-negative cases treated by surgical resection (2, 5, 6).

Table 1 *Prevalence of K-ras codon 12 mutations in lung tumors by diagnosis and detection method*

Diagnosis	No. of mutants/total (%)	
	PCR-PIREMA	ASO-H
Adenocarcinoma	13/28 (46)	6/28 (21)
Bronchoalveolar carcinoma	0/2 (0)	0/2 (0)
Large cell carcinoma	1/9 (11)	0/9 (0)
Squamous cell carcinoma	0/12 (0)	0/12 (0)
Carcinoid	0/2 (0)	0/2 (0)

The PCR-PIREMA assay is rapid, nonradioactive, and readily adaptable to processing large numbers of clinical samples and can also be used to detect all activating mutations in K- and N-*ras* codons 12, 13, and 61 (9). The protocol modifications reported here (decreased nucleotide and $MgCl_2$ concentrations in the PCR reactions) are critical to the assay specificity. When PCR-PIREMA is performed under "standard" PCR conditions (*i.e.*, 200 μM each nucleotide and 1.5–2.0 mM $MgCl_2$), the normal controls frequently fail to digest completely on enriched screening (Fig. 1, *Step 4*) even when a large excess of the screening enzyme is used (9). We reasoned that this difficulty might arise from the high misincorporation rate of *Taq* polymerase, of 1 error/10^4 bases under "standard" PCR conditions (19). The enrichment process would enrich not only for PCR products derived from authentic mutant alleles but also for PCR-introduced errors in codon 12, which would change the PCR product to *Bst*NI resistant. Standard PCR conditions maximize polymerization speed and yield, but at the expense of polymerase fidelity (19). It was empirically determined that 10 μM concentrations of each deoxynucleotide and 1.2 mM $MgCl_2$ in the first PCR reaction and 4 μM each deoxynucleotide and 0.6 mM $MgCl_2$ in the second reaction, with lengthened annealing and synthe-

sis times. optimized assay specificity while maintaining sufficient yield.

Because it requires minimal amounts of starting material, PCR-PIREMA can be easily applied to tissue biopsies or cells in body fluids prior to surgery; *e.g.*, we are using the method to detect mutations in bronchoalveolar lavage fluid from patients undergoing diagnostic bronchoscopy for suspected lung cancer (20). The clinical utility of *ras* as a biomarker for lung cancer has been suggested by investigators who found *ras* mutations in stored sputum samples from patients later diagnosed with lung adenocarcinoma (21). The assay used in that study, although sensitive, is labor intensive, necessitating cloning of sputum DNA followed by radioactive ASO-H. In contrast, the assay described here can be easily applied on a large scale in the clinical arena. With the demonstration that mutations in K-*ras* codon 12 are twice as prevalent as previously appreciated, their potential as clinically useful tumor markers is increased.

Acknowledgments

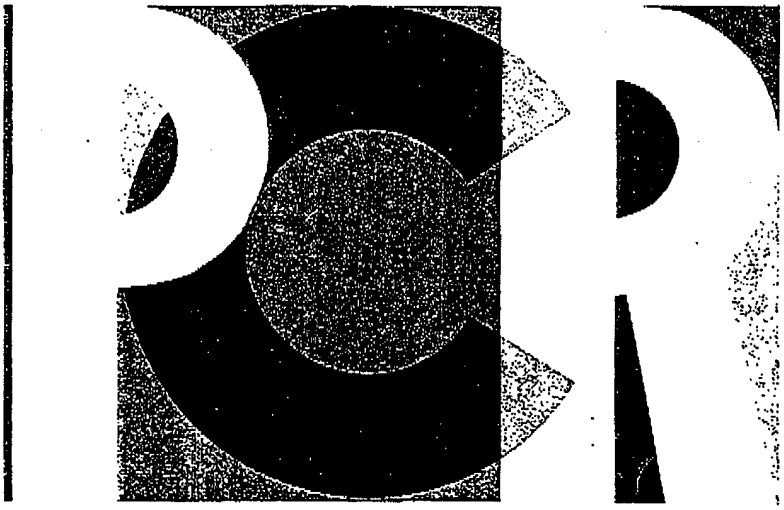
We thank Drs. Sjoerd Rodenhuis and Bert Top of the Netherlands Cancer Institute for supplying DNA samples and Dr. Rodenhuis for his careful review and helpful comments on the manuscript. We thank Drs. Jaishree Jagirdar, John Scholes, and Michael Ittman for assistance in procuring pathological samples and Dr. Joel Buxbaum for support.

References

1. Rodenhuis, S., and Slebos, R. J. C. Clinical significance of *ras* oncogene activation in human lung cancer. Cancer Res. (Suppl), *52:* 2665s–2669s, 1992.
2. Sugio, K., Ishida, T., Yokoyama, H., Inoue, T., Sugimachi, K., and Sasazuki, T. *ras* gene mutations as a prognostic marker in adenocarcinoma of the human lung without lymph node metastasis. Cancer Res., *52:* 2903–2906, 1992.
3. Mitsudomi, T., Viallet, J., Mulshine, J. L., Linnoila, R. I., Minna, J. D., and Gazdar, A. F. Mutations of *ras* genes distinguish a subset of non-small-cell lung cancer cell lines from small-cell lung cancer cell lines. Oncogene, *6:* 1353–1362, 1991.
4. Kobayashi, T., Tsuda, H., Noguchi, M., Hirohashi, S., Shimosato, Y., Goya, T., and Hayata, Y. Association of point mutation in c-Ki-*ras* oncogene in lung adenocarcinoma with particular reference to cytologic subtypes. Cancer (Phila.), *66:* 289–294, 1990.
5. Slebos, R. J. C., Kibbelaar, R. E., Dalesio, O., Kooistra, A., Stam, J., Meijer, C. J. L. M., Wagenaar, S. S., Vanderschueren, R. G. J. R. A., van Zandwijk, N., Mooi, W. J., Bos, J. L., and Rodenhuis, S. K-*ras* oncogene activation as a prognostic marker in adenocarcinoma of the lung. N. Engl. J. Med., *323:* 561–565, 1990.
6. Mitsudomi, T., Steinberg, S. M., Oie, H. K., Mulshine, J. L., Phelps, R., Viallet, J., Pass, H., Minna, J. D., and Gazdar, A. F. *ras* gene mutations in non-small cell lung cancers are associated with shortened survival irrespective of treatment intent. Cancer Res., *51:* 4999–5002, 1991.
7. Farr, C. J., Saiki, R. K., Erlich, H. A., McCormick, F., and Marshall, C. J. Analysis of *RAS* gene mutations in acute myeloid leukemia by polymerase chain reaction and oligonucleotide probes. Proc. Natl. Acad. Sci. USA, *85:* 1629–1633, 1988.
8. Ball, E. D., Mills, L. E., Neubauer, A., and Liu, E. Detection of minimal acute myeloid leukemia cells in bone marrow by probing for mutated *ras* genes using the polymerase chain reaction. Prog. Clin. Biol. Res., *333:* 499–506, 1990.
9. Jacobson, D. R., and Mills, N. E. A highly sensitive assay for mutant *ras* genes and its application to the study of presentation and relapse genotypes in acute leukemia. Oncogene, *9:* 553–563, 1994.
10. Wright, D. K., and Manos, M. M. Sample preparation from paraffin-embedded tissues. *In:* M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White (eds.), PCR Protocols: a Guide to Methods and Applications, pp. 153–158. San Diego, CA: Academic Press, Inc., 1990.
11. Ahuja, H. G., Foti, A., Bar-Eli, M., and Cline, M. J. The pattern of mutational involvement of *ras* genes in human hematologic malignancies determined by DNA amplification and direct sequencing. Blood, *75:* 1684–1690, 1990.
12. Toksoz, D., Farr, C. J., and Marshall, C. J. *ras* gene activation in a minor proportion of the blast population in acute myeloid leukemia. Oncogene, *1:* 409–413, 1987.
13. Syvänen, A. C., Söderlund, H., Laaksonen, E., Bengtström, M., Turunen, M., and Palotie, A. N-*ras* gene mutations in acute myeloid leukemia: accurate detection by solid-phase minisequencing. Int. J. Cancer, *50:* 713–718, 1992.
14. Bashey, A., Gill, R., Levi, S., Farr, C. J., Clutterbuck, J. L., Millar, J. L., Pragnell, I. B., and Marshall, C. J. Mutational activation of the N-*ras* oncogene assessed in primary clonogenic culture of acute myeloid leukemia (AML): implications for the role of N-*ras* mutation in AML pathogenesis. Blood, *79:* 981–989, 1992.
15. Rodenhuis, S., Slebos, R. J. C., Boot, A. J. M., Evers, S. G., Mooi, W. J., Wagenaar, S. S., van Bodegom, P. C., and Bos, J. L. Incidence and possible clinical significance of K-*ras* oncogene activation in adenocarcinoma of the human lung. Cancer Res., *48:* 5738–5741, 1988.
16. Birrer, M. J., and Brown, P. H. Application of molecular genetics to the early diagnosis and screening of lung cancer. Cancer Res., *52:* 2658s–2664s, 1992.
17. Szabo, E., Birrer, M. J., and Mulshine, J. L. Early detection of lung cancer. Semin. Oncol., *20:* 374–382, 1993.
18. Toksoz, D., Farr, C. J., and Marshall, C. J. *Ras* genes and acute myeloid leukemia. Br. J. Haematol., *71:* 1–6, 1989.
19. Eckert, K. A., and Kunkel, T. A. DNA polymerase fidelity and the polymerase chain reaction. PCR Methods Appl., *1:* 17–24, 1991.
20. Mills, N. E., Fishman, C. L., Scholes, J., Anderson, S. E., Rom, W. N., and Jacobson, D. R. Detection of K-ras oncogene mutations in bronchoalveolar lavage fluid for lung diagnosis. J. Natl. Cancer Inst., in press.
21. Mao, L., Hruban, R. H., Boyle, J. O., Tockman, M., and Sidransky, D. Detection of oncogene mutations in sputum precedes diagnosis of lung cancer. Cancer Res., *54:* 1634–1637, 1994.

EXHIBIT 7

AUGUST 1991
VOLUME 1 NUMBER 1



METHODS • AND • APPLICATIONS

The Ligase Chain Reaction

DNA Polymerase Fidelity

Vectorette PCR and Genomic Walking

Determining the 5′ End Sequence of mRNAs

Self-sustained Sequence Replication

Research

Optimization of the Polymerase Chain Reaction with Regard to Fidelity: Modified T7, *Taq*, and Vent DNA Polymerases

Lucy L. Ling, Phouthone Keohavong, Cremilda Dias, and William G. Thilly

Center for Environmental Health Sciences and Division of Toxicology, Whitaker College of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

The fidelity of DNA polymerases used in the polymerase chain reaction (PCR) can be influenced by many factors in the reaction mixture. To maximize the fidelity of DNA polymerases in the PCR, pH, concentrations of deoxynucleoside triphosphates, and magnesium ion were varied. Denaturing gradient gel electrophoresis was used to separate the polymerase-induced mutants from wild-type DNA sequences. Thermolabile modified T7 DNA polymerase, thermostable *Taq*, and Vent DNA polymerases were studied. Fidelity of all three DNA polymerases was sensitive to concentrations of deoxynucleoside triphosphates, magnesium ion, and pH. Within conditions that permitted efficient amplification, optimization with regard to these three factors yielded an average error rate in error/base pair incorporated of 7.2×10^{-5} for *Taq*, 4.5×10^{-5} for Vent, and 4.4×10^{-5} for modified T7 (Sequenase) DNA polymerases.

DNA amplification by the polymerase chain reaction (PCR)[1–3] can be accomplished by DNA polymerases from many sources.[1,2,4] Specific DNA polymerases have been reported to amplify with different efficiencies (i.e., yield of sequence amplification per cycle) and fidelities (frequency of polymerase-induced errors), with the kind and rate of error depending on the specific DNA polymerase and PCR conditions.[4–9] Each error, once initiated, is amplified along with the original wild-type sequences, increasing the fraction of polymerase-induced mutant sequences linearly with the number of amplification cycles. Under low-fidelity conditions, this mutant fraction can become significant: 40% of the amplification products are mutant when a 200-bp fragment is amplified one million-fold by a DNA polymerase with an error rate of 10^{-4},[4] such as Klenow fragment.[10] For some purposes such as sequencing a wild-type gene, the polymerase-induced mutations are generally distributed over the sequence of interest so that an accurate consensus sequence is usually obtained.[11,12] For studies of mutants within wild-type populations, it is vital that the polymerase-induced mutant sequences do not mask the preexisting mutant sequences.[11,13,14]

Due to its thermostability, *Taq* DNA polymerase has been the most widely used DNA polymerase. It does not have to be added every cycle, as was necessary for the thermolabile Klenow fragment referred to in the original descriptions of DNA amplification.[1,3,10] However, the fidelity of *Taq* has been reported to be 2×10^{-4} error/bp per duplication,[4,7–10] which renders it unsuitable for studies requiring both low noise and high amplification. Studies requiring a higher fidelity of amplification have used DNA polymerases such as modified T7 DNA polymerase (Sequenase) or T4 DNA polymerase,[4,13,14] which permitted amplification with error rates of about 3.4×10^{-5} and 0.3×10^{-5} error/bp per duplication, respectively.[4] However, both T4 and modified T7 DNA polymerases are thermolabile.

Recently, a new thermostable DNA polymerase called Vent DNA polymerase was isolated and made commercially available in a cloned form by New England Biolabs (Beverly, MA). Vent was reported to have a 3′→5′ proofreading exonuclease.[15] Our laboratory has been screening all available thermostable polymerases and has found that this new thermostable DNA polymerase, under conditions described for primer extension, permitted amplification with a useful degree of fidelity. Reaction conditions such as pH, concentrations of dNTP, and magnesium ion can greatly affect the fidelity of DNA polymerases.[5,7,9,16–20] Eckert and Kunkel used the M13mp2 fidelity assay and found that conditions of low pH with low and equimolar concentrations of dNTP to Mg^{2+} increased the fidelity of *Taq*.[7,9] We have ex-

1:63–69©1991 by Cold Spring Harbor Laboratory Press ISSN 1054-9803/91 $3.00

tended their observations using denaturing gradient gel electrophoresis (DGGE) to separate and enumerate all polymerase-induced mutant sequences from the correctly amplified sequences. To maximize the fidelity of *Taq*, Vent, and Sequenase, varied concentrations of dNTPs, pH, and magnesium were tested iteratively, and the optimal conditions with regard to these variables have been defined. We have confirmed the finding of Eckert and Kunkel[7,9] and defined conditions in which the *Taq* polymerase operates with a threefold improvement in fidelity during amplification in PCR relative to the most widely used reaction conditions.[4] More importantly, we find that the thermostable Vent polymerase under optimal conditions yielded a PCR amplification with both high fidelity and high efficiency which should be of general use to researchers using PCR.

MATERIALS AND METHODS

Materials

Taq polymerase was obtained from Perkin Elmer Cetus (Norwalk, CT). Vent polymerase, purified from *Thermococcus litoralis*, and its recombinant form cloned in *Escherichia coli* (designated Rec-Vent in this paper) were obtained from New England Biolabs (Beverly, MA). Sequenase (modified T7 DNA polymerase) was obtained from United States Biochemicals (Cleveland, OH). 2′-deoxynucleoside 5′-triphosphates (dNTPs) were obtained as 100 mM solution from Pharmacia (Piscataway, NJ). Primer P1, 5′-CATATATTAAATATACTCAC-3′; primer P2, 5′-TCCTGATTTTATTTC-TGTA-3′; and primer P4, 5′-GACTGA-ACGTCTTGCTCGAG-3′ were obtained from Synthetic Genetics (San Diego, CA). [γ-^{32}P]ATP (6000 Ci/mmole) was from New England Nuclear (Boston, MA) and T4 polynucleotide kinase was from New England Biolabs (Beverly, MA). The reagents in the 5′ DNA terminus labeling system were from Bethesda Research Lab (Gaithersburg, MD).

Procedure

The stock of DNA template for PCR was obtained by amplification from human genomic DNA, followed by purification with denaturing gradient gel to ensure a homogeneous population of the desired fragment of human HPRT exon 3 and flanking intron 2. This template was amplified from 10^5 copies to about 10^{12} copies and polyacrylamide gel purified. An aliquot (10^{10} copies) was further amplified 100-fold, under identical conditions with primer P4 end-labeled by T4 polynucleotide kinase.[13] The number of DNA copies generated was measured at different cycles to permit calculation of efficiency. Conditions that maintained exponential increases in DNA per cycle were used throughout this work. The ^{32}P-end-labeled DNA of the correct size was then purified from polyacrylamide gel, boiled, and reannealed so that each strand of any PCR-generated mutant homoduplexes could be hybridized to the complementary strand of the excess wild-type products to form heteroduplexes. These heteroduplexes were then separated from the wild-type homoduplexes by denaturing gradient gel electrophoresis. For each PCR condition, the total heteroduplex and the correctly amplified homoduplex fractions were measured for the calculation of fidelity in error/base pair incorporated per duplication.

PCR Procedure

pH, dNTP, and Mg^{2+} concentrations were experimental variables. $MgSO_4$ was used for *Taq* and Vent, while $MgCl_2$ was used for Sequenase. For Vent and *Taq*, the PCR reaction mixture was 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 0.1% Triton X-100, 0.1 μg/μl BSA, 1 μM each primer and 20 mM Tris buffer. Four units of enzyme were added for every 100 μl of reaction volume, with 3 more units of *Taq* polymerase added after 30 cycles. The amplification cycle consisted of 94°C for 1 min, 53°C for 2 min, and 70°C for 2 min. For Sequenase, the reaction mixture was 3 μM each primer, and 10 mM Tris buffer. The amplification cycle consisted of boiling for 1 min, cooling down to 37°C (1 minute), adding 1 unit (1 μl) of Sequenase, and incubation at 37°C for 2 min.[4]

Efficiency

To determine the efficiency of amplification at the exponential phase of DNA synthesis, a 5-μl aliquot of the PCR reaction mixture was removed and electrophoresed on 6% polyacrylamide gel together with a series of authentic standards whose numbers of molecules are known. The gel was stained with ethidium bromide and the PCR products of the correct size were measured by comparing the yield of the ultraviolet-induced fluorescence emitted by the ethidium bromide molecules intercalated into the DNA of the samples of DNA of the standard. The error of this method of quantitation is about twofold. However, this would affect the efficiency only slightly. For instance, a calculated efficiency of 70% due to 10^7-fold amplification after 30 PCR cycles, would be 75% if the DNA was actually twice as much. The efficiency of amplification was calculated from the formula:

$$(1 + Y)^n = \frac{\text{Final DNA copy number}}{\text{Initial DNA copy number}}$$

where Y is the efficiency per cycle and n is the number of PCR cycles under conditions of exponential amplification.[1,2,14]

Denaturing Gradient Gel Electrophoresis (DGGE)

To separate DNA duplexes containing only homologous wild-type sequences (WT homoduplexes) from the mutant/wild-type heteroduplexes, denaturing gradient gel electrophoresis was employed.[4] DNA was dissolved in a 30-μl solution of 400 mM NaCl, 10 mM Tris buffer at pH 7.5, 2 mM EDTA, boiled for 5 min, and reannealed for 5 hr at 65°C. The DNA was ethanol precipitated and separated on a denaturing gradient gel. The gel consists of a 12.5% polyacrylamide gel (acrylamide/bis-acrylamide = 37.5:1), containing a linearly increasing gradient of denaturant, parallel to the direction of electrophoresis, of 19–29% (vol/vol) where 100% denaturant is 7 M urea/40% formamide.[21] The gel was run for 15 hr at 150 volts while submerged in TAE buffer (40 mM Tris/20 mM acetic acid/2 mM EDTA, pH 8.3) at 60°C, dried, autoradiographed, and recorded on a phosphoImager screen (Molecular Dynamics, Sunnyvale, CA).

Determination of Fidelity

Fidelity of amplification was deter-

mined by analyzing the denaturing gradient gel exposed to the phosphoImager screen. The amount of excitation from the radioactive bands in the heteroduplex region (from the origin of the gel to the WT homoduplex band) and the band of WT homoduplex were measured. The background noise, such as nonspecific DNA binding in the gel, was measured from control lanes loaded with only DGGE purified wild-type DNA homoduplex sequences that had been subjected to the identical treatment of boiling, reannealing, and ethanol precipitation. The background control samples were run on the same denaturing gradient gels as the experimental samples. The heteroduplex fraction (HeF) due to polymerase amplification was calculated as: HeF = (Total heteroduplex counts – background heteroduplex counts)/total counts.

The fidelity was calculated[4] as:

$$f = \frac{\text{HeF}}{b\,d}$$

where f is the error rate (errors/base pair incorporated.duplication); HeF is heteroduplex fraction; b is the length of the single-strand low melting domain in which mutants can be detected (104 bp); and d is the number of DNA duplications (30).

RESULTS

Table 1 shows the values of pH, dNTPs, and magnesium with their corresponding efficiencies and fidelities that were observed in the studies with Sequenase, *Taq*, and Vent. Table 1A shows the efficiencies and error rates using conditions commonly used in PCR for *Taq* and Sequenase,[4] or in the case of Vent, those suggested by the supplier to be successful in primer extension. Table 1B shows the improvement brought about by the optimization effort.

Figure 1 is an autoradiogram of a denaturing gradient gel comparing the polymerization-induced mutations by Sequenase, *Taq*, and Vent. The band of wild-type homoduplex indicated by an arrow in the figure has been confirmed by sequencing. The positions of the mutant/WT type heteroduplexes of Vent are very similar to *Taq*, with a few exceptions. As conditions were adjusted for the improvement of fidelity, the positions of these heteroduplexes remained the same with changes only in their intensity relative to the WT homoduplex. The specific *Taq* and Sequenase point mutations corresponding to these bands have been previously reported.[4]

Optimization began with the conditions of Table 1A and iteratively examined the effects of dNTP concentration, pH, and then magnesium ion concentration. The effects of dNTP concentration on fidelity are shown in Figure 2. In a first round of experiments, the dNTP concentrations that permitted each polymerase efficiency to exceed 35% were defined. Within this range, *Taq* displayed some improvement in fidelity by either increasing or decreasing dNTP concentrations. The fidelity of Vent was somewhat improved by increasing dNTP as was that of Sequenase. Within the ranges shown on Figure 2, efficiencies were invariant for all polymerases, ~90% for Sequenase, ~70% for *Taq* and Vent. Based on these observations, 0.05 mM, 0.5 mM, and 3.5 mM dNTP for *Taq*, Vent, and Sequenase, respectively, were used in the next stage, study of pH effects.

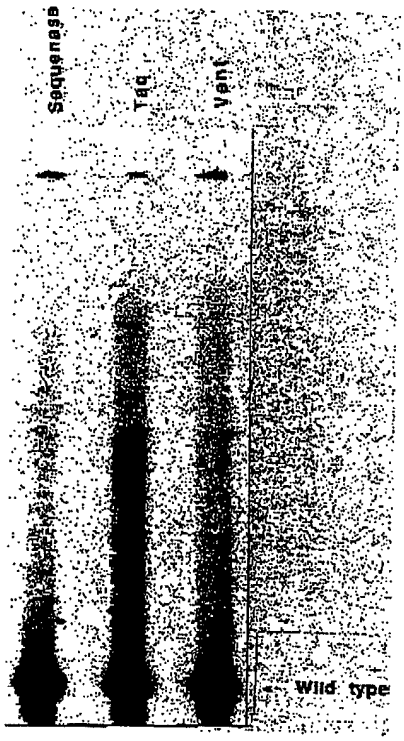


FIGURE 1 Analysis by DGGE of the PCR products after 10^9-fold amplification by Sequenase, *Taq*, and Vent DNA polymerases. The arrow points to the position where correctly amplified wild-type homoduplex focus.

The efficiencies of Vent and Sequenase were relatively invariant in the pH range of 7–9, but the efficiency of *Taq* decreased markedly below pH 8. As shown in Figure 3, within these ranges, *Taq* showed an improved fidelity when the pH was reduced to 8. Vent showed no change in fidelity from pH 8 to 9, but fidelity decreased at pH 7. The fidelity of Sequenase was invariant over the pH range of 7 to 9. Based on these observations, pH 8.0 was used for all three polymerases for the next step of the optimization procedure.

TABLE 1 Summary of the PCR Conditions

DNA polymerase	dNTP (mM)	pH	Mg (mM)	Efficiency (%)	Error rate (error/bp incorporated)
A. Initial conditions					
Sequenase	2.55	8.0	5	90	5.4×10^{-5}
Taq	0.2	8.8	2	70	2.0×10^{-4}
Vent	0.2	8.8	2	70	6.6×10^{-5}
B. Improved conditions					
Sequenase	3.5	8.0	2.5	90	4.4×10^{-5}
Taq	0.5–1.5	8.0	5	36	7.2×10^{-5}
Vent (and Rec-Vent)	0.5–1.5	8.5	7.5	70	4.5×10^{-5}

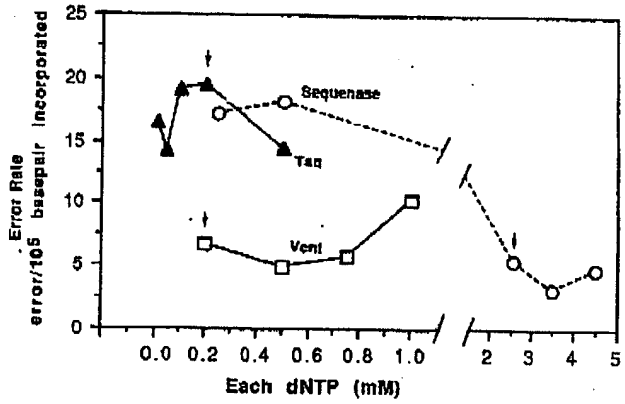


FIGURE 2 Fidelity of three DNA polymerases as a function of equimolar dNTP concentrations: (▲)*Taq* (pH 8.8, 2 mM $MgSO_4$), (□)Vent (pH 8.8, 2 mM $MgSO_4$), and (○)Sequenase (pH 8.0,5 mM $MgCl_2$). The efficiency was ~90% for Sequenase, ~70% for *Taq*, and 70% for Vent, for all dNTP concentrations shown. The arrows indicate the initial concentration of dNTP.

The next factor studied was the concentration of magnesium ion. Two separate batches of *Taq* (Perkin Elmer Cetus) were used. One batch of *Taq* showed acceptable amplification (~35%) only from 5 to 10 mM Mg. As shown in Figure 4, with the magnesium concentrations that permitted acceptably efficient amplification, the fidelity of *Taq* was improved both by increasing $MgSO_4$ to 10 mM or decreasing it to 0.5 mM. Vent polymerase and its bacterially produced cloned version, Rec-Vent, were found to have similar efficiency and fidelity. Their efficiencies were insensitive to $MgSO_4$ concentrations over the range studied, but below 2 mM $MgSO_4$, the fidelity of Vent and Rec-Vent DNA polymerase decreased drastically. A small but useful improvement in Sequenase fidelity occurred when $MgCl_2$ was decreased from 5 mM (Table 1A) to 2.5 mM.

At this point conditions had been defined which yielded fidelities about 4.5×10^{-5} for both Vent and Sequenase and as low as about 1.2×10^{-4} for *Taq* polymerase, each representing improvement of fidelity over that obtained using the original conditions of Table 1A.

To test for covariance of effects on fidelity among the variables studied, the lowest error rate conditions from Figure 4 were then probed in reverse order with regard to the effects of pH and then dNTP using the two thermostable polymerases, Vent (Rec-Vent) and *Taq* polymerases. As may be seen by comparing Figure 5 with Figure 3, the fidelity of *Taq* as a function of pH was not changed by the alterations in dNTP or $MgSO_4$ concentrations. The best fidelity of *Taq* was still obtained at pH 8 and efficiency dropped off below pH 8. However, the shift in dNTP and $MgSO_4$ concentrations did eliminate the increased error rate of Vent polymerase observed at pH 7 in Figure 3. Now a difference between native Vent and recombinant Vent polymerase was found: The efficiency of Rec-Vent dropped precipitously below pH 7.5.

Figure 6 shows the fidelity of *Taq*, Vent, and Rec-Vent polymerases as dNTP concentrations were varied at the pH and $MgSO_4$ concentrations that optimized fidelity. As seen by comparison with Figure 2, *Taq* was now found to amplify at dNTP concentrations above 0.5 mM, albeit with low efficiency (~36%). Remarkably, at pH 8 and 5 mM $MgSO_4$, the error rate of *Taq* dropped toward that seen previously for Vent and Sequenase, reaching a maximum fidelity at or above 0.5 mM dNTP. This effect was not found with the first round of optimizations (Figs. 2, 3, and 4), and emphasized the practical importance of reexamining the effect of the first two factors, namely pH and dNTP (Figs. 5 and 6).

The fidelity of both Vent and Rec-Vent was relatively invariant with dNTP concentration (Fig. 6), but once again a difference between Vent and Rec-Vent preparations was found in that only Vent showed efficient amplification below 0.5 mM dNTP. Because only Rec-Vent will be commercially available, except upon special request from the manufacturer, the Rec-Vent results should be noted in standard PCR applications.

DISCUSSION

The error frequency of different DNA polymerases has been measured by a variety of assays and ranged from 10^{-4} to 10^{-7} error/base pair incorporated.[4–7] In this study, a combination of PCR and DGGE provides a sensitive and reliable assay to measure the wide

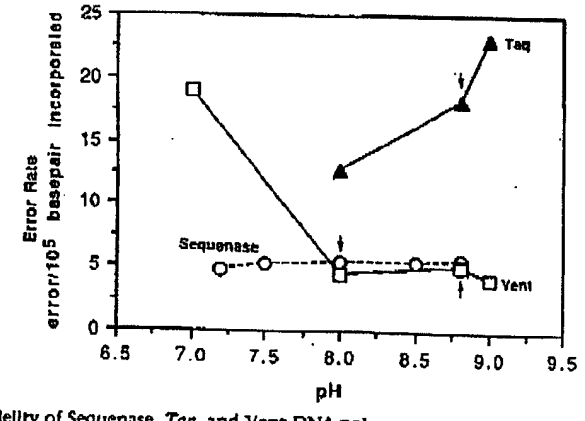


FIGURE 3 Fidelity of Sequenase, *Taq*, and Vent DNA polymerases as a function of pH: (▲)*Taq* (0.05 mM each dNTP, 2 mM $MgSO_4$), (□)Vent (0.5 mM each dNTP, 2 mM $MgSO_4$),(○) Sequenase (3.5 mM each dNTP, 5 mM $MgCl_2$). The efficiency was ~90% for Sequenase, ~70% for *Taq*, and Vent, except for Vent at pH 9 where it was ~45%. The arrows indicate the initial pH.

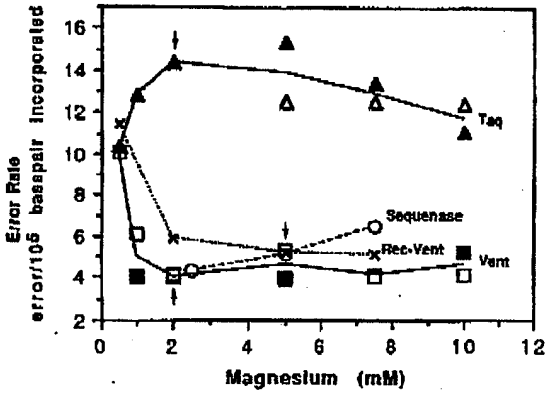


FIGURE 4 Fidelity of Sequenase, *Taq*, Vent, and Rec-Vent as a function of Mg^{2+} concentrations:(Δ ▲)*Taq* (0.05 mM each dNTP),(□ ■)Vent,(×)Rec-Vent (both 0.5 mM each dNTP), and (○)Sequenase (3.5 mM each dNTP). All reactions were at pH 8.0, using $MgCl_2$ for Sequenase and $MgSO_4$ for *Taq*, Vent, and Rec-Vent. The efficiency was 90% for Sequenase, ~70% for *Taq* Lot 3813(▲),except at ≤1 mM Mg^{2+} (36%), 40–63% for *Taq* Lot 3964(Δ),and ~70% for Rec-Vent and Vent, except at 0.5 mM Mg^{2+} (60%). The arrows indicate the initial Mg^{2+} concentration.

range of fidelity of different DNA polymerases over a variety of sequence contexts in the 104-bp target studied.[4] In its present form, error rates as low as 5×10^{-6} errors/base pair incorporated can be measured if care is taken to begin with a sufficiently large number of denaturing gradient gel-purified templates to eliminate jackpot effects. Another advantage is that DGGE provides a fast and accurate way to measure polymerase-induced mutants in virtually any sequence of interest to the investigators. All base-pair substitutions, small additions, and deletions are detected. The nature of the sequence changes can be determined by isolating and sequencing these DNA fragments.[4,13,21] However, if such information is not required, relative changes in the nature of polymerase-induced errors can be inferred from the pattern on the denaturing gradient gel alone.

The fidelity of DNA synthesis depends on the ability of DNA polymerases to discriminate against an incorrect nucleotide being ligated onto a primer-template, to extend from the mispair, and to correct the mispair by 3′→5′ exonucleolytic reversal of polymerization. This last ability is possible only for DNA polymerases with active 3′→5′ proofreading exonuclease. Among the three enzymes studied here, Vent is the only one known to have an active 3′→5′ exonuclease activity. However, its fidelity was matched by Sequenase and was 1.6-fold better than *Taq* polymerase, both of which are reported to be devoid of exonuclease activity.[8,22]

The fidelity of all three enzymes, Sequenase, *Taq*, and Vent, was shown to be sensitive to changes in pH, concentrations of dNTP, and Mg^{2+} ions. The fidelities of *Taq* and Sequenase were shown to have a different dependence on dNTP concentration than Vent. Previous studies of *Taq* had concluded that its fidelity is good at low equimolar concentration of total dNTP to Mg^{2+} concentration and that, in general, it is improved by lowering dNTP concentration, even in excess of Mg^{2+}.[7,9] Our results confirm these reports but also show that for both *Taq* and Sequenase, in conditions where the Mg^{2+} concentration is high and the total dNTP concentration is less than that of magnesium, fidelity is optimal at higher dNTP concentration. However, the fidelity of Vent, though invariant over a wide range of dNTP concentration, slightly decreased at high dNTP concentration. This may be due to diminished proofreading ability by high dNTP concentrations that has been observed in many DNA polymerases with 3′→5′ exonuclease activity such as *E. coli* Pol I and T4 polymerase.[6,17,23,24] High concentrations of dNTP are postulated to increase the rate of polymerization of the next correct nucleotide following the mispair, and thus decrease the probability of excision by the exonuclease.[23,25]

The effect of pH again showed no common trend on the fidelity of the three DNA polymerases tested. There was only a slight effect on the fidelity of Sequenase. The fidelity of Vent decreased at low pH only in conditions of low equimolar concentrations of dNTP to Mg^{2+}, whereas the fidelity of *Taq* progressively decreased as pH was raised. This supports the results of previous studies on Taq,[7,9] that at all Mg^{2+} and dNTP concentrations, the fidelity of *Taq* can be improved by

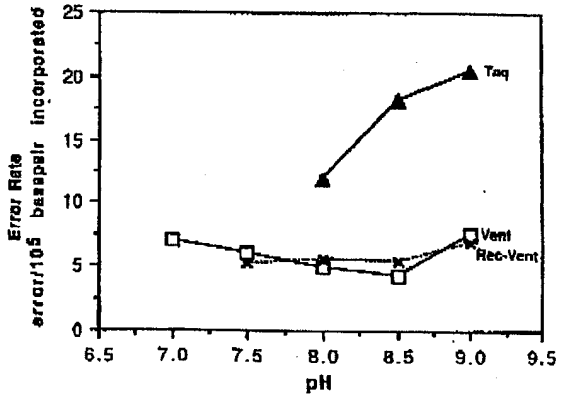


FIGURE 5 Fidelity of *Taq*, Vent, and Rec-Vent as a function of pH:(▲)Taq (0.05 mM each dNTP, 5 mM $MgSO_4$),(□)Vent and(×)Rec-Vent (both 0.5 mM each dNTP, 7.5 mM $MgSO_4$). The efficiency was ~40% for *Taq*, ~70% for Vent, and Rec-Vent (57% at pH 9).

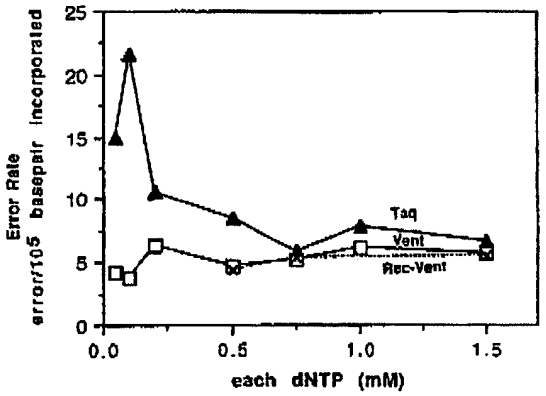


FIGURE 6 Fidelity of *Taq*, Vent, and Rec-Vent as a function of dNTP concentrations:(▲)*Taq* (5 mM $MgSO_4$, pH 8.0),(□)Vent, and(×)Rec-Vent (both 7.5 mM $MgSO_4$, pH 8.5). The efficiency was 34–38% for *Taq*, except at 0.5 and 1.5 mM dNTP (~59%), and ~70% for Vent and Rec-Vent.

lowering the pH. Mg^{2+} is used in this study as the activating divalent metal ion for DNA polymerization. It is considered a physiological activator because it is found in cells.[5] Depending on its concentration, Mg^{2+} has been shown to increase or decrease polymerase fidelity.[7,9,18,19] Fidelity has been shown to decrease by increasing the concentration of free Mg^{2+}.[7,9] Indeed, the fidelities of Sequenase and *Taq* were very slightly decreased by high concentrations of Mg^{2+}. The high fidelity of Vent was mainly invariant and decreased only at very low Mg^{2+} concentration. Other studies on DNA polymerases with 3′→5′ exonuclease activity such as *E. coli* Pol I have similarly noted that there was no effect of Mg^{2+} concentration on their fidelity,[18] except at inhibiting concentration.[19]

As adjustments in the PCR conditions were made, the intensities of all of the heteroduplexes relative to the WT homoduplex changed. However, there was no observable change in the bands of the heteroduplexes in the denaturing gradient gel. This indicates that hotspots for polymerase-induced errors were similar regardless of changes in the PCR conditions employed. The specific polymerase-induced mutants of Sequenase and *Taq* have been previously reported.[4] The positions of the heteroduplexes in the denaturing gradient gel are very similar between *Taq* and Vent, with a few exceptions. These bands have been isolated and are being sequenced.

The fidelity of Sequenase, *Taq*, and Vent can be usefully improved by appropriate changes in pH, concentrations of dNTP, and Mg^{2+} ion relative to conditions previously employed in PCR. It is probable that there are factors other than dNTP, pH, and Mg^{2+} that are also important for their fidelity in which overall mutant fractions of less than 10% are required for 100-bp sequences amplified a million-fold. The conditions represented in the results of this study were limited to those that allowed efficient amplification of a specific template, HPRT exon 3. It is probable that fidelity and efficiency are functions of the DNA sequence being amplified. The variations in the fidelity and efficiency of different lots of DNA polymerases is of practical importance in experimental procedure and should be noted. Thus, it is important to optimize the PCR conditions based on the DNA sequence of interest and the purpose of the study. The major improvement in the fidelity of *Taq* was achieved at very low amplification efficiency (~36%) whereas Vent (~70%) and Sequenase (~90%) consistently showed efficient amplification. Our results showed that Vent, a thermostable DNA polymerase, is as accurate in DNA synthesis as Sequenase and should be of use to researchers generating PCR products.

ACKNOWLEDGMENTS

This work was supported by grants from the U.S. Department of Energy (DE-FG02-86ER60448) and the National Institute of Environmental Health Sciences (P30-ES02109, 2P01-ES03926, 5P01-ES01640, and 5P42-ES04675).

REFERENCES

1. Saiki, R.K., S. Scharf, F. Faloona, K.B. Mullis, G.T. Horn, H.A. Erlich, and N. Arnheim. 1985. Enzymatic amplification of B-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230:** 1350–1354.
2. Mullis, K.B. and F.A. Faloona. 1987. Specific synthesis of DNA in vitro via a polymerase-catalysed chain reaction. *Methods Enzymol.* **155:** 335–350.
3. Kleppe, K., E. Ohtsuka, R. Kleppe, I. Molineux, and H.G. Khorana. 1971. Studies on polynucleotides. XCVI. Repair replications of short synthetic DNA's as catalyzed by DNA polymerases. *J. Mol. Biol.* **56:** 341–361.
4. Keohavong, P. and W.G. Thilly. 1989. Fidelity of DNA polymerases in DNA amplification. *Proc. Natl. Acad. Sci.* **86:** 9253–9257.
5. Loeb, L.A. and T.A. Kunkel. 1982. Fidelity of DNA synthesis. *Annu. Rev. Biochem.* **52:** 429–457.
6. Kunkel, T.A., L.A. Loeb, and M.F. Goodman. 1984. On the fidelity of DNA replication. *J. Biol. Chem.* **259:** 1539–1545.
7. Eckert, K.A. and T.A. Kunkel. 1990. High fidelity DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Nucleic Acids Res.* **18:** 3739–3744.
8. Tindall, K.R. and T.A. Kunkel. 1988. Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Biochemistry* **27:** 6008–6013.
9. Eckert, K.A. and T.A. Kunkel. 1991. The fidelity of DNA polymerases used in the PCR. In *Polymerase chain reaction: A practical approach* (eds. M.J. McPherson, P. Quirke, and G.R. Taylor), pp. 227–246. IRL Press at Oxford University Press, Oxford (in press).
10. Scharf, S.J., G.T. Horn, and H.A. Erlich. 1986. Direct cloning and sequence analysis of enzymatically amplified genomic sequences. *Science* **233:** 1076–1078.
11. Pääbo, S., D.M. Irwin, and A.C. Wilson. 1990. DNA damage promotes jumping between templates during enzymatic amplification. *J. Biol.*

Chem. **265:** 4718–4721.

12. Gyllensten, U. and H.A. Erlich. 1988. Generation of single stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA DQ alpha locus. *Proc. Natl. Acad. Sci.* **85:** 7652–7656.
13. Cariello, N.F., J.K. Scott, A.G. Kat, W.G. Thilly, and P. Keohavong. 1988. Resolution of a missense mutant in human genomic DNA by denaturing gradient gel electrophoresis and direct sequencing using in vitro DNA amplification: $HPRT_{Munich}$. *Am. J. Hum. Genet.* **42:** 726–734.
14. Keohavong, P., A.,G. Kat, N.F. Cariello, and W.G. Thilly. 1988. Laboratory methods: DNA amplification *in vitro* using T4 DNA polymerase. *DNA* **7:** 63–70.
15. Vent™ DNA polymerase technical bulletin from New England Biolabs. 1990 and 1991.
16. Goodman, M.F. 1988. DNA replication fidelity: Kinetics and thermodynamics. *Mutat. Res.* **200:** 11–20.
17. Kunkel, T.A., R.D. Sabatino, and R.A Bambara. 1987.Exonucleolytic proofreading by calf thymus DNA polymerase delta. *Proc. Natl. Acad. Sci.* **84:** 4865–4869.
18. Sirover, M.A., D.K. Dube, and L.A. Loeb. 1979. On the fidelity of DNA replication: Metal activation of Escherichia-Coli DNA polymerase. *J. Biol. Chem.* **254:** 107–111.
19. Kunkel, T.A. and L.A. Loeb. 1979. On the fidelity of DNA replication: Effect of divalent metal ion activators and deoxyribonucleoside triphosphate pools on in vitro mutagenesis. *J. Biol. Chem.* **254:** 5718–5725.
20. Weymouth, L.A. and L.A. Loeb. 1978. Mutagenesis during in vitro DNA synthesis. *Proc. Natl. Acad. Sci.* **75:** 1924–1928.
21. Myers, R.M., N. Lumelsky, L.S. Lerman, and T. Maniatis. 1985. Detection of single base substitution in total genomic DNA. *Nature* **313:** 495–498.
22. Huber, H.E., M. Russel, P. Model, and C.C. Richardson. 1986. Interaction of mutant thioredoxins of *Escherichia coli* with the gene 5 protein of phage T7. *J. Biol. Chem.* **261:** 15006–15012.
23. Kunkel, T.A. 1988. Exonucleolytic proofreading. *Cell* **53:** 837–840.
24. Kunkel, T.A. and A. Soni. 1988. Exonucleolytic proofreading enhances the fidelity of DNA synthesis by chick embryo DNA polymerase-γ. *J. Biol. Chem.* **263:** 4450–4459.
25. Ninio, J. 1975. Kinetic amplification of enzyme discrimination. *Biochimie* **57:** 587–595.

Received May 1, 1991; accepted in revised form July 2, 1991.

EXHIBIT 8

Avoiding false positives with PCR

S. Kwok and R. Higuchi

The exquisite sensitivity of the polymerase chain reaction means DNA contamination can ruin an entire experiment. Tidiness and adherence to a strict set of protocols can avoid disaster.

THE polymerase chain reaction (PCR)[1-3] is a powerful, exquisitely sensitive[3,4] technique with applications in many fields such as molecular biology, medical diagnostics, population genetics and forensic analysis. The use of specific DNA amplification, as reflected by the number of publications reporting the use of PCR, has indeed grown "exponentially" in recent months. We and others[5] are concerned, however, that some investigators may not be using adequate care in experimental design and execution when using PCR to detect only a few molecules of a target DNA sequence. A false positive or mistyping may occur when the majority of molecules to be detected arise from exogenous sources rather than from the sample itself.

Obviously, the fewer molecules one is trying to detect, the more one should guard against this possibility. The use of PCR for sensitive detection is complicated by the fact that the product of the amplification serves as the substrate for the generation of more product. A single PCR cycle produces very large numbers of amplifiable molecules that can potentially contaminate subsequent amplifications of the same target sequence. This kind of contamination has been termed PCR product "carryover" to differentiate it from contamination by naturally arising DNA (Fig. 1).

An analogy is useful to illustrate the scale of the contamination problem when using PCR to detect very few molecules. A typical PCR reaction can generate 10^{12} molecules of amplified DNA in a 0.1 ml reaction[3]. Imagine the uniform dilution of this number of molecules in a volume of liquid that fills an Olympic-sized swimming pool (~50 m × 25 m × 2 m). A 0.1 ml aliquot of liquid from this pool would contain 400 amplifiable molecules.

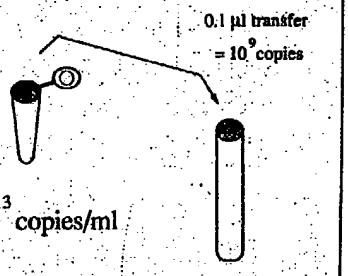


FIG. 1 The high concentration of amplifiable DNA in a completed PCR (typically ~ 10^{13} per ml) makes the inadvertent transfer of even a small volume to an unamplified sample quite serious. Reusing a pipette tip transfers 0.1 µl — roughly 10^9 copies of sequence. A microgram of human DNA contains only 1.4 × 10^5 copies of a single-copy gene.

To control contamination, one must prevent the physical transfer of DNA between amplified samples, and between positive and negative experimental controls. The following precautions have dramatically reduced false positive rates in our laboratories that perform PCR detection of HIV sequences[6] (in which as few as 15 copies of DNA per sample are routinely detected). Those who use PCR for less demanding procedures, such as the preparation of DNA fragments from plasmid DNA templates, may be able to adapt a less stringent form of these guidelines with successful results.

Tips for better PCR

Physically isolate PCR preparations and products. To prevent carryover of amplified DNA sequences, we prepare samples in a separate room or biosafety hood from that in which the reactions are performed. The UV germicidal lamps in most biosafety hoods quickly damage any DNA left on exposed surfaces, making it unsuitable for subsequent amplification (G. Sensabaugh, personal communication), and further eliminating the possibility of contamination between samples. Separate sets of supplies and pipetting devices are dedicated for sample preparation and for setting up reactions.

Autoclave solutions. We routinely autoclave deionized water and those buffer solutions used in PCR and sample preparation that can be autoclaved without affecting their performance. Autoclaving under conditions that provide bacterial decontamination degrades DNA to a very low molecular weight (N. Arnheim, personal communication). We also autoclave disposable pipette tips and microcentrifuge tubes. Primers, dNTPs and *Taq* DNA polymerase cannot be autoclaved.

Aliquot reagents. We divide reagents into aliquots to minimize the number of repeated samplings necessary. All reagents used in the PCR are prepared, divided and stored in an area that is free of PCR-amplified product. Similarly, oligonucleotides used for amplification are synthesized and purified in a PCR product-free environment. It is advisable to record the lot(s) of reagents used so that if contamination occurs, it can more easily be traced.

Use disposable gloves. Our researchers wear gloves and change them frequently, at the least when entering or reentering the isolation area. Changing gloves reduces the possibility of the transfer of amplifiable DNA from outside the isolation area and between samples.

Avoid splashes. Some types of sample tube have caps that require so much force to remove that liquid at the bottom of the tube may be splashed out. We use caps that do not require that much effort, and change gloves if such a splash occurs. It is also a good idea to spin any liquid down from the sides and top of the closed tube before attempting to open it.

Use positive displacement pipettes. The barrel of pipetting devices may become contaminated with aerosols containing sample DNA, leading to cross-contamination of samples. To prevent this, we use positive displacement pipettes with disposable tips and plungers.

"Premix" reagents. When possible, we mix reagents before dividing them into aliquots. All PCR reagents can be combined into a "premixture" which can then be pipetted into reaction vessels containing DNA. This minimizes the number of sample transfers and the chances for sporadic contamination. When dispensing the mixture, we pipette a "no DNA" negative control last so that it reflects the total reagent handled.

Add DNA last. Non-sample components, such as mineral oil, pre-mixed dNTPs, primers, buffer and enzyme, should be added to the amplification vessels before sample DNA. This minimizes cross-contamination by reducing the number of opportunities for the inadvertent transfer of DNA. After the addition of DNA, we cap each tube before proceeding to the next sample.

Choose positive and negative controls carefully. We do not use a highly concentrated solution of plasmid DNA containing the target sequence as our positive control, because it would introduce as many amplifiable molecules into the sample preparation area as a typical PCR. If plasmid DNA containing the target sequence is used as a positive control, it should be diluted substantially. Depending upon the detection system, as few as 100 copies of target may suffice as a positive control.

We include "no DNA" reagent controls and negative sample controls with each set of amplifications. The reagent controls should contain all the necessary components for PCR, except template DNA.

Negative sample controls should not contain target sequences, but should have gone through all the sample preparation steps. Choosing negative controls for our HIV studies was complicated by the fact that PCR enabled HIV sequences to be detected from samples which were negative by all other tests. In the end, we used samples from low-risk individuals with well-known histories.

Avoiding other pitfalls

The amount of PCR product generated is sometimes insufficient, requiring re-amplification after enrichment by gel electrophoresis. If the amount of amplified target sequence DNA in a particular gel slice is very low, one should be wary of cross-contamination from analogous PCR products or plasmid DNA containing the target sequence run in other lanes of the gel. To prevent carryover from equipment used in previous experiments, gel apparatus and combs should be soaked in 1 M HCl to depurinate any residual DNA, new razor blades should be used to excise each gel band, and the surface of the UV transilluminator should be covered with a fresh sheet of plastic wrap for each gel. Other potential sources of contamination include purified restriction fragments of target sequence, dot-blot apparatus, microtome blades, centrifuges, centrifugal vacuum devices and dry ice or ethanol baths. Most of these items are also amenable to treatment, if necessary, with 1 M HCl.

A final caution

If there is any doubt at all about a critical result, it is best to repeat the experiment again. The negative controls described above can rule out reagent contamination, but cannot guarantee against sporadic contamination events. Fortunately, the odds of a sporadic contamination event occurring twice the same way are very low. The net error rate of a series of tests is also usually lower than that of a single trial, even if the sporadic error rate of the repeated test is higher. PCR is simple and rapid, and consumes so little of most samples that repeat experiments can be performed, even in forensic work[7].

Kits for performing PCR are available from Perkin-Elmer Cetus under the trademark GeneAmp. □

Shirley Kwok and Russell Higuchi are at the Departments of Infectious Diseases and Human Genetics, Cetus Corporation, Emeryville, California 94608, USA. For more information, fill in reader service number 100.

1. Erlich, H.A., Gelfand, D.H. & Saiki, R.K. *Nature* **331**, 461–462 (1988).
2. Saiki, R.K. *et al. Science* **239**, 487–491 (1988).
3. Faloona, F. & Mullis, K. *Meth. Enzym.* **155**, 335–350 (1987).
4. Li, H. *et al. Nature* **335**, 414–417 (1988).
5. Lo, Y.-M.D., Mehal, W.Z. & Fleming, K.A. *Lancet* **2**, 679 (1988).
6. Ou, C. *et al. Science* **239**, 295–297 (1988).
7. Higuchi, R., von Beroldingen, C.H., Sensabaugh, G.F. & Erlich, H.A. *Nature* **332**, 543–546 (1988).

BioExpo on show

At next week's BioExpo in Paris, France, new products range from an automated loop inoculation device to a crystallized Tris buffer preparation.

THE MAbTrap G from Pharmacia LKB is designed for the **purification of monoclonal and polyclonal IgG** from ascites fluid, serum and cell culture supernatants (*Reader Service No. 101*). Because Protein G binds only IgG and its subclasses, separating out other immunoglobulins, Pharmacia LKB says the MAbTrap G offers high capacity, high flow rate and high recovery.

Tired of laboriously flaming loops when streaking petri plates? The Swiss company Tecnomara has an **automatic flame inoculation** system, dubbed the Rondoflame (*Reader Service No. 102*). The base unit of the system is an automatic bunsen burner fireboy that can hold four inoculating loops inserted into specialized holders on a rotation assembly at 90° angles to the plane of rotation. One loop is always in the optimal position for flaming: the fireboy burns only as long as necessary to flame the loop and automatically advances the loop to the take-up position. The device can even be programmed to rotate clockwise or counterclockwise for left- and right-handed use. The Rondoflame is marketed at a cost of $1,350 (US).

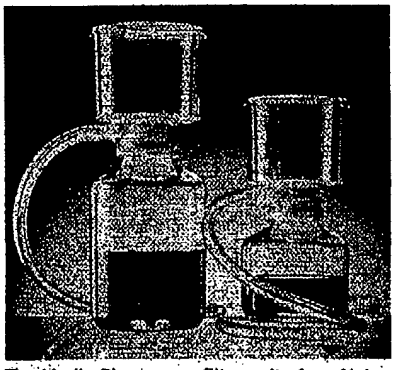
Boehringer Mannheim is now offering its Ultrapure **Tris buffer in crystalline form** (*Reader Service No. 103*). The company says the buffer dissolves easier in solution than powdered buffers without clumping, and that it is free of contaminating proteases, DNases and RNases. At a purity of over 99.9 per cent, BMB says the Tris buffer is totally clear in solution, low in metal content and costs $160 (US) for 5 kg.

A new range of supports for **ion exchange chromatography** is being added to the product list at IBF Biotechnics (*Reader Service No. 104*). The Trisacryl Plus supports, based on the Trisacryl matrix, are of fundamental benefit, says IBF Biotechnics: they are resistant to extremes in pH, high flow rates, freezing, autoclaving and bacterial contamination. The Trisacryl Plus range is available in 300-ml bottles for the M grade (40–80 μm bead size) for $118 (US) and in 1-litre bottles for the LS grade (80–160 μm) for $355 (US).

Trisacryl Plus chromatography media.

Beckman Instruments has a new series of **HPLC autosamplers** that operate on-line as part of the company's System Gold series of liquid chromatographs or stand alone as part of another system (*Reader Service No. 105*). The model 507 autosampler is designed for laboratories that analyse large sample volumes on a daily basis. Operating features include a four-quadrant, refrigerated sample holder with a 96-vial capacity and a separate vial for rinse liquid. Vial access is direct, and column temperature control is available. The $10,500-16,225 (US) model 507 also offers variable injection volumes, column switching, and pre-column reagent addition, mixing and reaction. □

These notes are compiled by Cary Prince from information provided by the manufacturers. To obtain additional information about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal, and apply only within the country indicated.



The Media-Plus serum filter units from Nalge now come with 90-mm nylon membranes.

market. The nylon membranes have a 0.2-mm pore size, and are designed for use in tissue culture applications with sensitive cell lines, or in situations in which a wetting agent may cause problems. The nylon Media-Plus filter units are radiation-sterilized and cost $192 (US) for a case of 12.

Three types of CelPrep **gel entrapment media** are now available from FMC Bioproducts (*Reader Service No. 113*). FMC Bioproducts says its CelPrep Alginate, CelPrep Agarose and CelPrep Carrageenan provide different cellular environments and enhance tissue culture growth. Gel entrapment offers a three-dimensional cell growth matrix which can be used with a variety of cell types. The company tests each medium to ensure low endotoxin levels and freedom from mycoplasma contamination. Each gel medium is available in 125-ml quantities for $30 (US).

To create a two-compartment system for transport studies, co- culture, chemotaxis and invasion assays, Costar recommends its Transwell **cell culture chamber insert** (*Reader Service No. 114*). The inserts provide a microporous basis for cell attachment, and are available with either polycarbonate or collagen-treated membranes. Polycarbonate membranes are available in pore sizes of 0.1 μm, 0.4 μm, 3.0 μm, 5.0 μm and 8.0 μm, with diameters to fit 24-well and 6-well plates. The collagen-treated, transparent membranes offer increased visualization of cell outlines during phase contrast microscopy, says Costar. The collagen-treated membranes cost $131 (US) for 24-well plates and for $143 for 6-well plates.

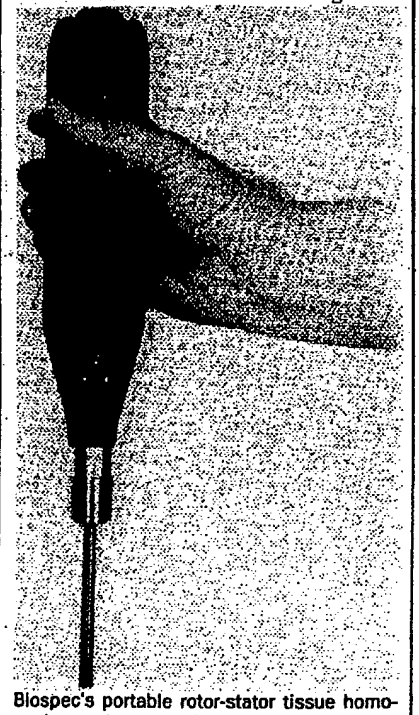
Correction

THE product review "Avoiding false positives with PCR" (*Nature* 339, 237; 18 May, 1989) was unfortunately published without the authors' final corrections and contains errors of emphasis and fact. First, the authors did not fully agree with the editor's summary that "tidiness and adherence to a strict set of procedures could prevent disaster", but preferred the idea that common sense and an awareness of these procedures could help prevent mistakes. Second, although the number of potentially contaminating molecules produced by PCR is large, the authors had intended to point out that there is an analogous situation in routine microbiology, where such standard operating procedures allow one to cope. Third, the authors intended to emphasize that controlling contamination entails thinking about the *relative* numbers of target molecules in samples handled together and the consequences of inadvertent physical transfer between them. Finally, it was to have been suggested that one should use only the number of PCR cycles necessary to achieve desired sensitivity and avoid the amplification in negative controls of inconsequential levels of contamination.

Other persisting errors include the number of molecules left in 0.1 ml of a dilution into an Olympic-sized swimming pool (40 not 400); the number of copies of a single-copy gene in one μg of human DNA (3×10^3); and reference and typographical errors. Also, it was not intended to suggest that UV lights be used in a biosafety hood *while* manipulating samples, but only between uses of the hood. Reprints of the article in corrected form are available from the authors. □

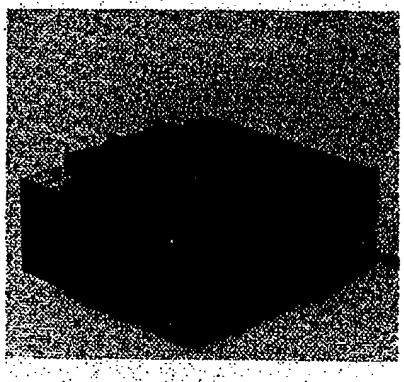
Sliced and diced

Biospec is introducing a **portable rotor-stator homogenizer** that can disperse, homogenize and emulsify samples of volumes between 0.5 ml and 50 ml (*Reader Service No. 115*). Suspended material is drawn into the homogenizer probe by a screw-shaped rotor turning at high speeds. The suspension is then repeatedly recycled through six close-tolerance slits at a rate of 20,000 times per minute. Biospec says the rotor-stator homogenizer disintegrates and disperses most materials in one minute or less. The cordless, two-speed, hand-held instrument is rechargeable in three hours. Its probe is four inches in length and 0.25 inches in diameter.



Biospec's portable rotor-stator tissue homogenizer recharges in three hours.

A slicer for finely cutting biopsy tissues with less cellular damage than homogenizers is being introduced by Balzers (*Reader Service No. 116*). The Balzers **tissue slicer** is used for mechanically cutting tissues originating from biopsy or very small organs for subsequent embedding or ultramicrotomy. The instrument is capable of sectioning tissue as thin as 15-20 μm, says Balzers. The sectioning thickness of the slicer can be set in microns using its calibrated micrometer head, a property Balzers says is useful for antibody labelling and peroxidase studies. The $2,860 (US) bench-top unit can also make sections, cubes and wedges of tissue.



The tissue slicer from Balzers also makes sections, cubes and wedges.

Cambridge Instruments is now selling the Reichert-Jung Cryocut 1800 **cryostat**, which features a stainless steel microtome (*Reader Service No. 117*). The microtome is incorporated into an ergonomically designed freezer cabinet and has a spindle feed assembly and roller-bearing slideway that can cut specimens in 1-μm sections. The Cryocut 1800's refrigeration unit provides finger-touch temperature control, and its motor-driven course feed responds with button control. The availability of both automatic and manual 12-minute defrost cycles offer frost-free operation. □

These notes are compiled by Cary Prince from information provided by the manufacturers. To obtain additional information about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal, and apply only within the country indicated.

EXHIBIT 9

Translational Research and Epithelial Carcinogenesis: Molecular Diagnostic Assays Now—Molecular Screening Assays Soon?

*Michael J. Birrer**

The application of modern molecular biologic techniques to the complex health problem of human cancer has yielded impressive results in expanding our understanding of the basic scientific mechanisms underlying tumor development. Despite these gains, the clinical impact of these insights has remained elusive. Nowhere has this difficulty been clearer than in the attempted translation of recent basic science discoveries into clinical applications in the diagnosis and detection of epithelial cancers. It is equally clear, however, that a new age is dawning in which important clinical goals based on basic science principles will be achievable (*1*). In this issue of the Journal, Mills et. al. (*2*) report one such example in which modern molecular biologic techniques have been used to identify molecular markers in clinically relevant human tissue specimens with important diagnostic and predictive value.

Mills et al. detected mutations in codon 12 of the K-ras oncogene in bronchoalveolar lavage fluid from 86 patients with a clinical suspicion of lung cancer. Of 52 patients with confirmed lung cancer, 16 (31%) had activated ras mutations in their bronchoalveolar lavage fluid. Of note, 14 of these fluids were cytologically negative. All mutated ras genes characterized in lavage fluids were identical to those found in the corresponding tumor (if available). Mills et al. point out that no activated ras genes were detected in any bronchoalveolar lavage fluids from patients with a diagnosis other than non–small-cell lung cancer (not including four positive lavages from patients with a strong clinical suspicion of lung cancer but without a tissue diagnosis of lung cancer). This result produced an assay with 31% sensitivity (16 of 50 lung cancer cases) and with 88% specificity (30 of 34) for detecting lung cancer.

This article by Mills et al. is another major step toward creating molecular-based diagnosis/detection assays for epithelial cancers. Mutated genes have been detected in various human fluids in cell debris that is sloughed off from tumors (*3*). For instance, activated ras genes have been identified in stool specimens from patients with colorectal cancer (*4*); mutated p53 genes have been found in urine samples from patients with bladder cancer (*5*); and both genes have been identified in sputum specimens from patients who eventually developed lung cancer (*6*). These landmark studies firmly established that genetic lesions within tumors could be identified in the bodily secretions or fluids that bathe these tumors. These initial studies, however, retrospectively analyzed small numbers of stored specimens and frequently analyzed tumor specimens first and then searched for the identical molecular marker in the matched draining material. Thus, they suffer from certain statistical limitations and selection bias, which do not allow us to accurately ascertain the potential clinical impact of these assays. The study by Mills et al. addresses some of these problems by assaying a much larger number of patients in a prospective fashion. This sample size makes the sensitivity and specificity results more convincing. Unfortunately, the study design is not optimal, since it is not entirely clear that the authors were blinded to the clinical diagnosis or to the molecular status of the tumor at the time of the analysis of the bronchoalveolar lavage fluid. This approach may have biased their analysis and ultimately the results. Nevertheless, this study does bring us one step closer to an effective molecular detection assay for lung cancer.

In general, there are two major classes of problems associated with the development of these types of assays. First, there are technical obstacles that need to be overcome to develop an assay that is both sensitive and specific at detecting the marker of interest. Techniques to detect activated ras genes in tumor specimens have been available for several years and have undergone continual evolution. Standard assays include NIH3T3 cell transformation (*7*) and ribonuclease (RNase) protection assays (*8*) but each was quickly recognized as labor intensive and relatively insensitive (RNase protection) to apply to large numbers of specimens. A second generation of polymerase chain reaction (PCR)-based assays include PCR amplification with allele-specific oligonucleotide hybridization (PCR–ASO-H) (*9*), PCR amplification using mismatched primers creating a "designed restriction fragment length polymorphism (RFLP)" (*10*), and PCR with single-strand conformation polymorphism (PCR–SSCP) analysis (*11*). These techniques are applicable to large numbers of specimens and are, in general, more sensitive than previous assays. However, they still require that 5%-20% of the specimen contain an activated ras gene for detection (*12*). Further modification includes the cloning of the PCR product and allele-specific oligonucleotide hybridization, which improved sensitivity and provided quantification (*4*). This particular technique has identified mutated ras genes in stool and sputum from patients with cancer. More recent improvements in these assays have involved enrichment of the PCR amplification schema involving "designed RFLP," which has increased overall sensitivity while maintaining the general applicability of the assay (*13*). Mills et al. (*2*) use a version of this type of assay (primer-introduced restriction with enrichment for mutant alleles [PCR–

**Correspondence to:* Michael J. Birrer, M.D., Ph.D., National Institutes of Health, Key West Research Center, 9610 Medical Center Dr., Rm. 300, Rockville, MD 20850.

PIREMA]), employing cycles of PCR and restriction enzyme digestion to enrich and detect activated ras genes. This assay can be optimized to detect one activated ras gene among 10^6 ras alleles and has already been demonstrated to identify mutated ras genes in adenocarcinomas of the lung which were previously scored as normal by PCR–ASO-H (*14*). Improvements in sensitivity are critical given the molecular heterogeneity of cancer. It is conceivable that only a small proportion of many tumors or preneoplastic lesions may contain activated ras genes. In addition, diagnostic specimens, such as bodily secretions or fluids, are extensively contaminated with normal cells; this contamination would require additional sensitivity in a test to detect the presence of mutated ras genes.

Equally important is the specificity of the test. In general, false-positive results with activated ras gene assays have not been a problem (*2,4,6*). As the tests become increasingly sensitive, involving more cycles of PCR amplification and enrichment, however, the possibility of artifactual mutation becomes more likely. Since Taq polymerase makes one error per 10^4 bases, Taq-induced mutations are a concern (*15*). Most investigators take great care to optimize PCR conditions to lower the Taq polymerase error rate (*16*). In addition, the PCR–PIREMA technique has a number of specific controls, primer selections, and verification steps that help to ensure assay specificity (*12*).

The second problem in the development of molecular screening/detection assays is our lack of understanding of the precise role of the gene or marker of interest in the development of cancer. Critical to the use of a molecular marker for the diagnosis or detection of cancer is its frequency of occurrence within the tumor versus other disease states. A gene mutation that is rarely present in a tumor would not be an effective target for a sensitive detection assay, while a genetic lesion frequently present in non-tumor disease states would produce diagnostic assays that had low specificity. For lung cancer, ras mutations have been detected in approximately 17% of tumors (24% of adenocarcinomas) by PCR–ASO-H (*17*) and in up to 32% of tumors (48% of adenocarcinomas) by PCR–PIREMA (*14*). These results are confirmed by the present study in which 16 (31%) of 52 lung cancer cases were found to have activated ras genes. With this kind of frequency of occurrence, it is clear that a detection assay for lung cancer based on mutated ras genes will have limitations in sensitivity. It should be noted, however, that highly specific but insensitive assays can be linked, producing a test that screens for a battery of markers usually resulting in an improvement in overall sensitivity. Thus, if specific detection tests for other genetic lesions identified in lung cancer, such as p53 mutations (*18*) or 3p deletions (*19*), could be developed and combined with the assay used by Mills et al., the result would likely be an assay with an increased sensitivity for detecting lung cancer. In fact, this technology could be combined with other non-molecular detection assays for lung cancer such as those that employ monoclonal antibodies (*20*) in an attempt to optimize sensitivity and specificity.

Critical to the development of molecular detection assays ultimately is their clinical utility. One such application might be as a diagnostic adjunct that could be used to simplify diagnostic algorithms, morbidity, and/or cost. Present detection assays for activated ras genes in lung cancer appear to be close to formal testing for this purpose. For instance, the diagnosis of lung cancer may be made in the appropriate clinical setting (strong clinical suspicion of lung cancer) by PCR–PIREMA identification of K-ras in bronchoalveolar lavage fluid alone without requiring further invasive tests, such as thin-needle aspirations, transbronchial biopsy, or thoracotomy (unless otherwise clinically indicated for staging or treatment). Of course, it would still involve the invasive procedure of bronchoaveolar lavage, but this procedure results in a considerably lower morbidity than the other procedures listed. In addition, it is conceivable that even this procedure might be eliminated if the ras detection assay could be effectively applied to sputum, as has been suggested already (*6*).

A more important application is the use of these assays to reduce lung cancer morbidity and mortality through early detection. Lung cancer is frequently advanced at diagnosis and is rarely curable at such a stage (*21*). Therefore, any approach that would allow for the early detection of this disease at a stage in which it might be surgically curable could have important clinical consequences. Unfortunately, multiple attempts to establish an effective lung cancer-screening assay using the modalities of chest x-ray and/or sputum cytology has thus far failed (*22*). Therefore, there is considerable interest in applying the above-discussed technology to this problem. Unfortunately, this is a much more demanding application for molecular assays and, at present, more development and testing are certainly required to ascertain if this application will be possible. For instance, in the article by Mills et al. (*2*), we are not told the precise stage of the patients' tumors, only that all of these patients had clear radiographic evidence of cancer. This suggests that these patients did not have early (and therefore curable) disease. Therefore, the results should be interpreted as strongly supportive of a detection assay for lung cancer but not necessarily for the early detection of this disease. To be an effective early detection assay, it would have to identify the presence of malignancy well before it becomes clinically evident (radiographically or otherwise). This situation requires that the target marker be present during early tumor development. For oncogene mutations, this implies an early mutational event such as a tumor initiation event. Unfortunately, for carcinogen-driven tumors such as lung cancer, these events may all too frequently be an occurrence within the affected field (field cancerization) (*23,24*). Thus, assays that are sensitive enough to detect all initiation events in a given target gene may lead to detection assays that are relatively nonspecific. It may be necessary to tailor the sensitivity of the assay or to select a specific marker(s) to identify a clone of cells that would be more predictive of the development of cancer. For lung cancer, these markers are not yet known. Mutations of p53 and deletions of 3p have been identified in preneoplastic lesions of the lung but whether they occur only in cells destined for progression to malignancy remains unknown (*24,25*). From this angle, ras mutations may actually turn out to be a good marker for lung cancer, as there is evidence that they do not occur in preneoplastic lesions but do occur in carcinoma in situ and early invasive cancers (*26,27*). Thus, one would be identifying a relatively early clone of cells that will later produce a lung malignancy. Obviously, substantial work will be required to identify the correct molecular marker(s) and assay(s)

that will yield a sensitive and specific molecular-screening assay for lung cancer. Ultimately, this process will require careful validation by prospective screening trials involving patients at high risk for lung cancer with lung cancer morbidity and mortality as end points.

In conclusion, we are clearly witnessing a revolution in our understanding of the causes of and molecular mechanisms involved in cancer development. The clinical implications of this revolution are just now presenting themselves. With continued identification of the specific molecular events occurring during the development of epithelial cancers, such as lung cancer and with the further development of molecular technology, accurate screening assays for a wide range of epithelial cancers will undoubtedly become available.

References

(1) Birrer MJ, Brown PH: Application of molecular genetics to the early diagnosis and screening of lung cancer. Cancer Res 52(9 Suppl):2658s-2664s, 1992

(2) Mills NE, Fishman CL, Scholes J, et al: Detection of K-ras oncogene mutations in bronchoalveolar lavage fluid for lung cancer diagnosis. J Natl Cancer Inst 87:1056-1060, 1995 MILLS E

(3) Sidransky D: Identifcation of gene mutations in cancer screening. *In* Early Detection of Cancer (Srivastava S, Lippman S, Hong W, et al., eds). Armonk: Futura Publishing Co, 1994

(4) Sidransky D, Tokino T, Hamilton SR, et al: Identification of ras oncogene mutations in the stool of patients with curable colorectal tumors. Science 256:102-105, 1992

(5) Sidransky D, Von Eschenbach A, Tsai YC, et al: Identification of p53 gene mutations in bladder cancers and urine samples. Science 252:706-709, 1991

(6) Mao L, Hruban RH, Boyle JO, et al: Detection of oncogene mutations in sputum precedes diagnosis of lung cancer. Cancer Res 54:1634-1637, 1994

(7) Liu ET: The role of ras gene mutations in myeloproliferative disorders. Clin Lab Med 10:797-807, 1990

(8) Winter E, Yamamoto F, Almoguera C, et al: A method to detect and characterize point mutations in transcribed gene: amplification and overexpression of the mutant c-Ki-ras allele in human tumor cells. Proc Natl Acad Sci U S A 82:7575-7579, 1985

(9) Sugio K, Ishida T, Yokoyama H, et al: ras gene mutations as a prognostic marker in adenocarcinoma of the human lung without lymph node metastasis. Cancer Res 52:2903-2906, 1992

(10) Kumar R, Dunn LL: Designed diagnostic restriction fragment length polymorphisms for the detection of point mutations in ras oncogenes. Oncogene Res 4:235-241, 1989

(11) Suzuki Y, Orita M, Shiraishi H, et al: Detection of ras gene mutations in human lung cancers by single-strand conformation polymorphism analysis of polymerase chain reaction products. Oncogene 5:1037-1043, 1990

(12) Jacobson DR, Mills NE: A highly sensitive assay for mutant ras genes and its application to the study of presentation and relapse genotypes in acute leukemia. Oncogene 9:553-563, 1994

(13) Teneriello MG, Ebina M, Linnoila RI, et al: p53 and Ki-ras gene mutations in epithelial ovarian neoplasms. Cancer Res 53:3103-3108, 1993

(14) Mills NE, Fishman CL, Rom WN, et al: Increased prevalance of K-ras oncogene mutations in lung adenocarcinoma. Cancer Res 55:1444-1447, 1995 MILLS I

(15) Eckert KA, Kunkel TA: DNA polymerase fidelity and the polymerase chain reaction. PCR Methods Appl 1:17-24, 1991

(16) Kwok S, Higuchi R: Avoiding false positives with PCR [published erratum appears in Nature 339:490, 1989]. Nature 339:237-238, 1989

(17) Slebos RJ, Kibbelaar RE, Dalesio O, et al: K-ras oncogene activation as a prognostic marker in adenocarcinoma of the lung. N Engl J Med 323:561-565, 1990

(18) Takahashi T, Nau MM, Chiba I, et al: p53: a frequent target for genetic abnormalities in lung cancer. Science 246:491-494, 1989

(19) Brauch H, Johnson B, Hovis J, et al: Molecular analysis of the short arm of chromosome 3 in small-cell and non-small-cell carcinoma of the lung. N Engl J Med 317:1109-1113, 1987

(20) Tockman MS, Gupta PK, Myers JD, et al: Sensitive and specific monoclonal antibody recognition of human lung cancer antigen on preserved sputum cells: a new approach to early lung cancer detection. J Clin Oncol 6:1685-1693, 1988

(21) Ginsberg R, Kris M, Armstrong J: Cancer of the lung. *In* Cancer: Principles and Practice of Oncology (DeVita VT Jr, Hellman S, Rosenberg SA, eds). Philadelphia: Lippincott, 1989

(22) Richert-Boe KE, Humphrey LL: Screening for cancers of the lung and colon [*see* comment citation in Medline]. Arch Intern Med 152:2398-2404, 1992

(23) Slaughter DP, Southwick HW, et al: "Field cancerization" in oral stratification of squamous epithelium: clinical implications of multicentric origin. Cancer 6:963-968, 1953

(24) Sozzi G, Miozzo M, Pastorina U, et al: Genetic evidence for an independent origin of multiple preneoplastic and neoplastic lung lesions. Cancer Res 55:135-140, 1995

(25) Bennett WP, Colby TV, Travis WD, et al: p53 protein accumulates frequently in early bronchial neoplasia. Cancer Res 53:4817-4822, 1993

(26) Sugio K, Kishimoto Y, Virmani AK, et al: K-ras mutations are a relatively late event in the pathogenesis of lung carcinomas. Cancer Res 54:5811-5815, 1994

(27) Li ZH, Zheng J, Weiss LM, et al: c-k-ras and p53 mutations occur very early in adenocarcinoma of the lung. Am J Pathol 144:303-309, 1994

EXHIBIT 10

ARTICLES

Detection of K-ras Oncogene Mutations in Bronchoalveolar Lavage Fluid for Lung Cancer Diagnosis

*Nancy E. Mills, Charles L. Fishman, John Scholes, Sibyl E. Anderson, William N. Rom, Daniel R. Jacobson**

Background: Lung cancer is the leading cause of cancer deaths in the United States. A long-standing goal of cancer researchers has been to develop tests that would facilitate earlier diagnosis and treatment of lung cancer and thereby decrease mortality from this disease. Because cancer results from the accumulation of a variety of genetic events (e.g., mutations, rearrangements, and deletions) in genes controlling cell growth and differentiation, these changes might serve as diagnostically useful molecular markers. Activation of the K-ras oncogene by point mutations in codon 12, which occurs in many cases of lung adenocarcinoma, may serve as one such clinically useful molecular marker. For detection of K-ras point mutations in bronchoalveolar lavage fluid, in which small numbers of malignant cells are mixed with a population of predominantly genetically normal cells, the sensitivity of commonly used assays for ras mutations risks false-negative results. *Purpose:* By applying a highly sensitive assay, we investigated whether detection of K-ras codon 12 mutations in samples of bronchoalveolar lavage fluid could be clinically useful in diagnosing lung cancer. *Methods:* We developed a highly sensitive assay for detecting K-ras codon 12 mutations based on an enriched polymerase chain reaction (PCR) technique. This technique was applied to 87 specimens of bronchoalveolar lavage fluid specimens that were obtained from 86 patients, and associated tumor biopsy specimens obtained from 35 of these patients who underwent diagnostic bronchoscopy for clinically suspected lung cancer. Lavage fluid specimens were also obtained from nine patients undergoing nondiagnostic bronchoscopy. Statistical comparisons were performed by using the two-tailed Fisher's exact test. *Results:* Of 52 patients with confirmed lung cancer, samples of bronchoalveolar lavage fluid from 16 patients contained K-ras codon 12 mutations, including 14 (56%) of 25 patients with lung adenocarcinomas, one (33%) of three with bronchoalveolar carcinomas, one (20%) of five with large-cell carcinomas, and none of the 14 with squamous cell carcinomas. Mutations were detected in four additional cases in which cancer was suspected but had not been histologically confirmed. Tissue samples from 35 of the patients all yielded the identical K-ras codon 12 genotype found in the corresponding samples of bronchoalveolar lavage fluid. No mutation was found in any sample from 30 patients with diagnoses other than non–small-cell lung cancer. Thus, for those cases in which tissue was available and tested, the sensitivity and specificity of detecting K-ras mutations in bronchoalveolar lavage fluid for diagnosing K-ras mutation-positive lung cancer were both 100%. For nine patients, K-ras mutations were detected in bronchoalveolar lavage fluid obtained during otherwise nondiagnostic bronchoscopies. *Conclusions:* Our data demonstrate that sensitive detection of K-ras codon 12 mutations can serve as an important adjunct to cytology in the diagnosis of lung cancer. *Implications:* Detection of these mutations could lead to earlier cancer diagnosis and less need for invasive diagnostic procedures. [J Natl Cancer Inst 87:1056-1060, 1995]

Lung cancer is the leading cause of cancer deaths in the United States (*1*). A long-standing goal of cancer researchers has been to develop tests that would facilitate earlier diagnosis and treatment of lung cancer and thereby decrease the mortality from this disease (*2*). Because cancer results from the accumula-

**Affiliations of authors:* N. E. Mills, Department of Medicine, Divisions of Hematology and Oncology, New York University School of Medicine, and Kaplan Cancer Center, New York University Medical Center, NY; C. L. Fishman, Department of Medicine, Division of Pulmonary and Critical Care Medicine, New York University School of Medicine; J. Scholes, Department of Pathology, New York University School of Medicine; S. E. Anderson, New York University School of Medicine; W. N. Rom, Department of Medicine, Division of Pulmonary and Critical Care Medicine, New York University School of Medicine, and Kaplan Cancer Center, New York University Medical Center; D. R. Jacobson, Department of Medicine, Division of Hematology, New York University School of Medicine, Kaplan Cancer Center, New York University Medical Center, and Research and Medical Services, New York Veterans Affairs Medical Center.

Correspondence to: Daniel R. Jacobson, M.D., Research Service 151, New York Veterans Affairs Medical Center, 423 E. 23 St., New York, NY 10010.

See "Notes" section following "References."

tion of a variety of genetic events (e.g., mutations, rearrangements, and deletions) in genes controlling cell growth and differentiation, these changes could serve as diagnostically useful molecular markers, leading to earlier treatment and potentially improving the outcome of patients with lung cancer (*3,4*).

Patients being evaluated for suspected lung cancer often undergo fiberoptic bronchoscopic biopsy to obtain a tissue diagnosis. During bronchoscopy, bronchoalveolar lavage with saline injected through the bronchoscope is also performed. When, because of its location, a lesion is inaccessible to bronchoscopic biopsy, or the biopsy obtained is nondiagnostic, a diagnosis of cancer is occasionally made by cytologic examination of the bronchoalveolar lavage fluid, but this is much less sensitive than biopsy for diagnosis (*5*). Body fluids sometimes contain cells or cell debris bearing the same oncogene mutations characterizing their related tumors, as has been shown for ras mutations in stool from patients with colorectal tumors (*6*) and for p53 (also known as TP53) mutations in urine from patients with bladder cancer (*7*). Similarly, mutations associated with lung cancer might be detectable in bronchoalveolar lavage fluid, thus providing clinically useful lung cancer markers and improving the diagnostic yield of bronchoscopy.

One such potential biomarker is the K-ras oncogene. The three ras genes (N-ras, K-ras, and H-ras) encode signal-transduction proteins. Point mutations (nearly always in codons 12, 13, and 61) confer transforming properties on the ras genes by yielding proteins with constitutive activation that generate continuous signals to proliferate (*8*). In the largest study (*9*) of ras mutations in human lung cancer, K-ras mutations in codon 12 predominated, and they were found in 45 (17%) of 258 non–small-cell lung cancer (NSCLC) samples obtained at surgical resection, primarily in adenocarcinomas (43 [24%] of 181 samples). In addition, patients whose tumors had ras mutations had a significantly poorer outcome than patients with tumors of the same stage but without ras mutations. We have recently demonstrated that the prevalence of ras mutation originally reported in 24% of patients with lung adenocarcinomas was an underestimate and that the true prevalence in these patients was about 50% (*10*).

Cancer cells in bronchoalveolar lavage fluid are always mixed with large numbers of genetically normal, nucleated cells, including alveolar macrophages, white blood cells, and normal bronchial epithelial cells (*11*). Moreover, in any given tumor, only a fraction of the malignant cells may contain a ras mutation, as in acute myelogenous leukemia (*12-15*); thus, detection of ras mutations in bronchoalveolar lavage fluid requires a sensitive assay. In most recent studies, assays for ras mutations have utilized the polymerase chain reaction (PCR) to amplify ras exon 1 (containing codons 12 and 13) and exon 2 (containing codon 61), followed by allele-specific oligonucleotide hybridization. This method requires that, to be detectable, a ras mutation must be present in at least 10% of the background ras alleles (*16-18*). In studies of clinical specimens, such as bronchoalveolar lavage fluid, in which small numbers of mutation-containing tumor cells are mixed with larger numbers of genetically normal cells, the insensitivity of PCR/allele-specific oligonucleotide hybridization risks false-negative results.

We have developed a more sensitive assay for ras mutations in our laboratory, which we have termed "polymerase chain reaction–primer-introduced restriction with enrichment for mutant alleles" (PCR–PIREMA) (*10,15*). Using this assay, we investigated whether detection of K-ras codon 12 mutations in samples obtained at bronchoscopy, particularly bronchoalveolar lavage fluid, could be clinically useful in diagnosing lung cancer.

Materials and Methods

DNA was isolated from bronchoalveolar lavage fluid samples obtained during 87 clinically indicated (as determined by the patient's primary physicians), bronchoscopic procedures performed on 86 patients with radiographically suspected lung cancer (and also from nine patients undergoing nondiagnostic bronchoscopy) at the New York University Medical Center. Lavage was performed after wedging of the bronchoscope in an airway adjacent to the radiographically defined lesion under fluoroscopic guidance. By the procedure, 100 mL of normal saline was introduced in 20-mL aliquots and was recovered after 15 seconds, with a 60-mL recovery volume. For cytologic examination, the bronchoalveolar lavage fluid was preserved in 50% ethanol and centrifuged, and the pellet was fixed on a glass slide. For the ras assay, samples of the bronchoalveolar lavage fluid were processed according to standard protocols developed for purification of DNA: incubation in buffer containing proteinase K and detergents, followed by organic extraction (*19*).

Paraffin-embedded samples of biopsy and resection tissues from these patients were obtained when available, and they were deparaffinized (*20*) prior to DNA isolation as described earlier. All DNA samples were desalted and concentrated to 5-10 μL in Microcon or Centricon 100 concentrators (Amicon Inc., Beverly, MA) twice prior to PCR.

The PCR–PIREMA assay for mutations in K-ras codon 12 was performed as described (*10*). The original assay detected as little as one mutant allele in N-ras codon 12 or 61 per 10^6 normal alleles (*15*). As modified and simplified in the present study, the assay detects mutant alleles present at the level of 0.1% (*10*). Briefly, PCR around K-ras codon 12 was performed by using a mismatched primer that introduced a *Bst*NI restriction site into PCR products derived from normal, but not mutant, alleles. *Bst*NI digestion of the PCR products left only PCR products derived from mutant alleles intact ("enrichment"), after which further PCR selectively amplified the mutant PCR products (analyzed by *Bst*NI digestion and gel electrophoresis; "enrichment screening"). Additional PCR was performed on the enriched products by using a panel of other mismatched primers, which introduced new restriction sites for various enzymes, dependent on the specific mutation present in the original sample. These PCR products were then digested with the appropriate enzymes, electrophoresed on agarose gels, and stained with ethidium bromide ("verification"). Digested bands seen on verification identified mutations present in the starting material.

All samples were subjected to the entire PCR–PIREMA process at least twice. Extensive measures were taken to prevent cross-contamination of samples (*15*). Multiple normal control samples and negative control samples (no DNA in the PCR reaction) were included in all experiments. PCR–PIREMA results have been previously shown to agree with results obtained by PCR/allele-specific oligonucleotide hybridization and DNA sequencing but at an increased sensitivity (*15*).

All *P* values were calculated from two-tailed tests of statistical significance (Fisher's exact test). These investigations were approved by the New York University Medical Center Institutional Review Board in accord with an assurance filed with and approved by the U.S. Department of Health and Human Services.

Results

K-ras codon 12 mutations were found in samples of bronchoalveolar lavage fluid from 20 patients: 14 (56%) of 25 patients with lung adenocarcinoma, one (33%) of three patients with bronchoalveolar carcinoma, one (20%) of five patients with large-cell carcinoma, and four patients with pulmonary lesions highly suspicious for cancer both radiographically and clinically, but who remain undiagnosed (Table 1; Fig. 1). Nineteen of

Table 1. Frequency of K-ras codon 12 mutations in samples of bronchoalveolar lavage fluid by diagnosis*

Diagnosis	No. of patients	Mutant (%)
Adenocarcinoma	25	14 (56)
Bronchoalveolar carcinoma	3	1 (33)
Large-cell carcinoma	5	1 (20)
Squamous cell carcinoma	14	0 (0)
NSCLC (unable to specify)	5	0 (0)
All confirmed lung cancer histologies	52	16 (31)
Cancer diagnosis suspected but not proven histologically	4	4
Diagnosis other than NSCLC†	30	0 (0)
Total No. of diagnoses	86	20

*Numbers and percentages refer to numbers of patients, not numbers of bronchoalveolar lavage fluids. One patient (Table 2, patient 3) had two lavage fluid samples analyzed, each of which was positive; this patient is counted only once here.

†The 30 patients with diagnoses other than NSCLC had the following pathologic diagnoses: pneumonitis (nine patients), nonspecific inflammation (nine patients), no pathologic diagnosis (five patients), inflammatory atypia (four patients), metastatic sarcoma (one patient), Kaposi's sarcoma (one patient), and sarcoidosis (one patient).

the 21 samples of mutation-positive bronchoalveolar lavage fluids (including two samples from one patient) were cytologically negative for malignancy. No mutation was found in 14 samples of bronchoalveolar lavage fluid from patients with squamous cell carcinoma. Moreover, no mutation was found in any sample from 30 patients with diagnoses other than NSCLC (Table 1). For the 20 patients with lavage fluid positive for K-ras codon 12 mutations, the predominant mutations were the same as those that have been predominant in other reports (*9,21,22*) of K-ras codon 12 mutations in lung cancer: TGT (seven patients), GAT (six patients), and GTT (five patients). Two patients were found to have less common mutations: CGT (one patient) and AGT (one patient). Among the 52 patients with confirmed lung cancer, the percentage of mutation-positive samples was significantly higher for patients with adenocarcinoma than for patients with either squamous cell carcinoma (Fisher's exact test; $P = .00035$) or all other histologies combined (Fisher's exact test; $P = .0002$), confirming previous observations that ras mutations are most common in adenocarcinoma.

Paraffinized tissue samples were available from 16 of the 20 patients with mutation-positive, bronchoalveolar lavage fluid (bronchoscopic biopsy specimens in 12 patients, resected tumor specimens in two patients, both a transthoracic needle aspiration and a resected sample in one patient, and both a transbronchial biopsy specimen and autopsy tissue in one patient). In all cases, the corresponding tissue and bronchoalveolar lavage fluid contained the identical ras mutation. Tissue samples were available for analysis from 19 of the 36 patients with confirmed lung cancer but normal bronchoalveolar lavage fluid ras genotypes, and none contained a K-ras codon 12 mutation. Thus, although not all tumors could be tested, for those cases in which tissue was available and tested, the sensitivity and specificity of ras mutation detection in bronchoalveolar lavage fluid for diagnosing ras mutation-positive lung cancer were both 100%.

Ten samples of bronchoalveolar lavage fluid that contained ras mutations were obtained (from nine patients) during entirely

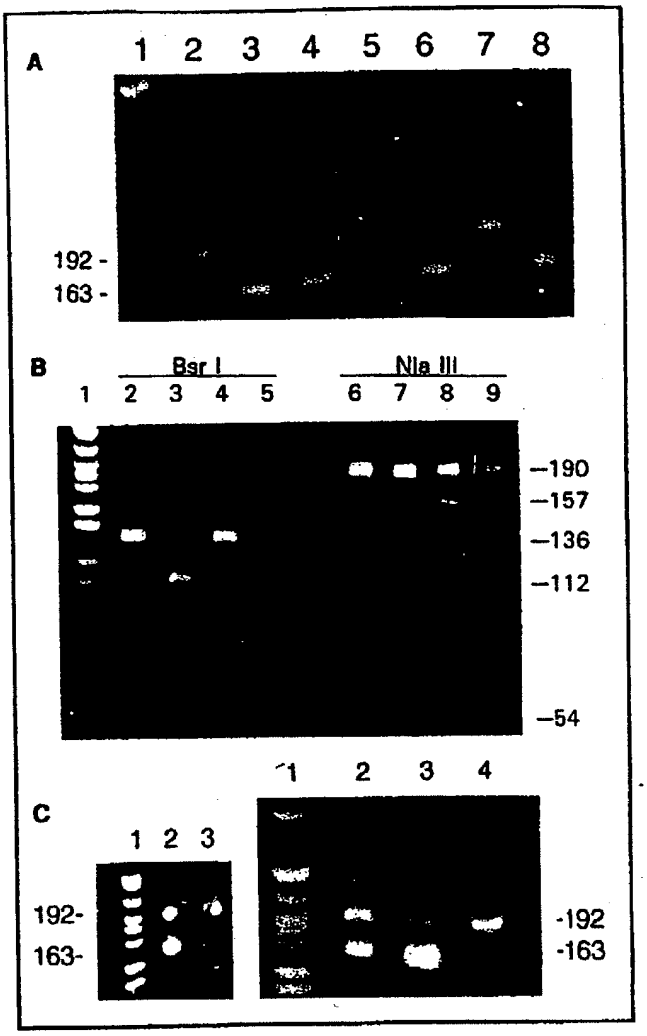


Fig. 1. Analysis of bronchoalveolar lavage fluids and associated tissue for K-ras codon 12 mutations by PCR–PIREMA. A) Enriched screening (*Bst*NI digestion), where a digestion-resistant band indicates that a mutation is present but does not identify the specific mutation. Lane 1 = size marker (upper band) and undigested PCR product of 192 base pair (bp) (lower band). Lanes 2 and 7 = bronchoalveolar lavage fluid and autopsy tissue, respectively, from patient 3 (Table 2), each revealing a mutation (digestion-resistant PCR product of 192 bp) as well as the normal allele (163 bp). Enriched screening of DNA isolated from bronchoalveolar lavage fluid obtained at this patient's second bronchoscopy also revealed a mutation, as did DNA isolated from her second transbronchial biopsy (which was histologically nondiagnostic), and all four samples verified to the mutant codon GTT, as in panel B, lane 3. Lane 5 = bronchoalveolar lavage fluid from another patient, demonstrating a mutation. Lanes 3, 4, and 6 = bronchoalveolar lavage fluid samples revealing a normal pattern (complete digestion). Lane 8 = normal control. B) and C) Verification, where a digested band indicates the specific mutation present. B) Verification for GTT mutation by *Bsr* I digestion and for TGT mutation by *Nla* III digestion. Lane 1: size marker. Lanes 2-4 and 6-8 = bronchoalveolar lavage fluid samples digested with *Bsr* I and *Nla* III, respectively. Lanes 5 and 9 = normal control digested with *Bsr* I and *Nla* III, respectively. The 112-bp band in lane 3 indicates a GTT mutation, and the 157-bp band in lane 8 indicates a TGT mutation. The bands denoting the normal-sequence PCR product in the *Bsr* I digests (136 bp) are smaller than the bands denoting the normal-sequence PCR product in the *Nla* III digests (190 bp) because of an invariate *Bsr* I site elsewhere in the PCR product. C) Verification for AGT mutation by *Bfa* I digestion (left panel) and for GAT mutation by *Bsp*HI digestion (right panel). Left panel, lane 1 = size marker. Lanes 2 and 3: *Bfa* I digest of a bronchoalveolar lavage fluid sample and normal control, respectively. The 163-bp band in lane 2 indicates an AGT mutation. Right panel, lane 1: size marker. Lanes 2 and 3 = bronchoalveolar lavage fluid samples digested with *Bsp*HI. Lane 4 = normal control. The 163-bp band in lanes 2 and 3 indicates GAT mutations.

nondiagnostic bronchoscopic procedures, i.e., transbronchial biopsy, brushings, and lavage, which did not yield a histologic cancer diagnosis (Table 2). In five patients (Table 2, patients 1-5), the diagnosis of lung cancer was made after further invasive diagnostic procedures. Cancer was not diagnosed in one individual (patient 3) until after death; samples of bronchoalveolar lavage fluid obtained both 1 year and 6 months prior to the patient's death contained the same mutation, and a nondiagnostic transbronchial biopsy specimen obtained at the second bronchoscopy was also positive for this mutation, as was tumor tissue obtained at autopsy (Fig. 1, A).

Four of the patients with ras mutation-positive bronchoalveolar lavage fluid obtained at completely nondiagnostic bronchoscopy (Table 2, patients 6-9) still lack a histologic cancer diagnosis despite strong clinical and radiographic suspicion of lung cancer. Patient 9 had a large mass in the left lower lobe of the lung and died before histologic diagnosis could be obtained (no postmortem examination was performed). Patient 6 had a large mass in the right lower lobe with mediastinal adenopathy, but this individual had multiple medical problems and refused further diagnostic intervention. Patient 7 had a resected lung adenocarcinoma 2 years previously and presented with a hilar mass, and both bronchoscopy and transthoracic needle aspiration were nondiagnostic. Patient 8 had a right upper lobe coin lesion that was unresponsive to antituberculous therapy. Further evaluation was needed to determine these patients' histologic diagnoses.

Discussion

Our data demonstrate that the ras mutation is a clinically useful biomarker for lung cancer and that sensitive ras mutation detection can serve as an important adjunct to bronchoscopy with bronchoalveolar lavage in the diagnosis of lung cancer. To our knowledge, this is the first report of oncogene detection in bronchoalveolar lavage fluid. We have detected ras mutations in 21 samples of bronchoalveolar lavage fluid from patients with proven or suspected lung cancer: 19 of these samples were cytologically negative. In nine patients, the bronchoalveolar lavage fluid cytology and corresponding bronchoscopic biopsy and brushings all failed to yield a histologic cancer diagnosis. Furthermore, no ras mutation was seen in any patient with a diagnosis other than NSCLC, confirming both the specificity of this test for malignancy and its utility as a marker for lung cancer. For some patients, such as patients 1-5 shown in Table 2, the clinical use of this assay would have meant earlier lung cancer diagnosis and avoidance of additional invasive diagnostic procedures. For patient 3, who was not diagnosed with lung cancer until after death, this assay could have provided a diagnosis during the patient's lifetime and offered the possibility of appropriate therapy.

Patients 6-9 (Table 2), with ras mutation-positive samples of bronchoalveolar lavage fluid obtained at completely nondiagnostic bronchoscopy, still lack a histologic cancer diagnosis despite strong clinical and radiographic suspicion of lung cancer. For at least one individual (patient 9), histologic proof of cancer will never be obtained. For patients 6 and 7, a histologic cancer diagnosis is still being pursued. If these patients are eventually diagnosed with lung cancer, as is predicted both clinically and by their ras mutation status, then the use of this assay might have been of value in achieving earlier lung cancer diagnosis for them. At present, we do not know for certain that a ras mutation-positive sample of bronchoalveolar lavage fluid indicates that a patient definitely has lung cancer, though we believe that this is likely to be the case. At the least, we suggest that the finding of a mutation in samples of bronchoalveolar lavage fluid indicates that an aggressive effort to obtain a histologic diagnosis should be pursued.

Previously, the use of mutant ras genes as a molecular marker for lung cancer has been limited by both the relatively low frequency with which such mutations have been reported to occur and by the fact that methods of mutation detection have been up to this point relatively insensitive or cumbersome. The prevalence of 56% ras mutation positivity in this study in both bronchoalveolar lavage fluid and tissue from patients with lung adenocarcinoma (the subtype most frequently associated with ras mutations) is more than twice that which has been reported in several large studies (*9,16,22,23*) of ras activation in lung cancer. We have recently found that the prevalence of ras mutations in resected lung cancer samples is greater when analyzed by PCR–PIREMA (about 50%) than when the same samples are analyzed by PCR/allele-specific oligonucleotide hybridization (*10*). Thus, the use of a more sensitive assay results in a higher estimation of ras mutation frequency in lung cancer. The clinical utility of ras as a biomarker for lung cancer has also been suggested by investigators (*24*) who found a prevalence of ras

Table 2. K-ras codon 12 mutations in samples of bronchoalveolar lavage fluid obtained at nondiagnostic bronchoscopy

Patient No.*	Mutant codon	Diagnosis	Procedure yielding diagnosis	Tissue with mutation
1	TGT	Large-cell carcinoma	Transthoracic needle aspiration	Transbronchial biopsy
2	GTT	Adenocarcinoma	Transthoracic needle aspiration	Transthoracic needle aspiration resection
3a	GTT	Bronchoalveolar carcinoma	Autopsy	Autopsy and second endobronchial biopsy
3b	GTT			
4	TGT	Adenocarcinoma	Resection	Resection
5	GAT	Adenocarcinoma	Biopsy of adrenal metastasis	Endobronchial biopsy
6	GAT	Pending	None	Endobronchial biopsy
7	CGT	Pending	None	No biopsy
8	TGT	Pending	None	Not available
9	TGT	Died without tissue diagnosis	None	No biopsy

*Samples 3a and 3b were derived from two bronchoscopies done 6 months apart on the same patient. The histologic diagnosis is undetermined for patients 6-9, but lung cancer is suspected clinically. No biopsy was done on patients 7 and 9; no remaining pathologic tissue is available from patient 8.

mutations comparable with that reported in the present study in stored sputum samples from a small group of patients later diagnosed with lung adenocarcinoma. Although sensitive, the assay used in that study is labor-intensive, necessitating cloning of sputum DNA followed by radioactive allele-specific oligonucleotide hybridization. In contrast, the PCR–PIREMA assay can be easily applied on a large scale, so it has potential use in the clinical arena, as demonstrated here.

Because many lung cancers do not harbor K-ras codon 12 mutations, testing bronchoalveolar lavage fluid for this mutation is not of use in all cases. Since clonal ras mutations are thought to be highly specific for the diagnosis of cancer, however, they offer benefit as a diagnostic tool in cases where they are positive. Furthermore, as our assay detects K-ras codon 12 mutations in about 50% of lung adenocarcinoma cases, this test should be useful in more patients than would have been expected on the basis of several previous reports (*9,16,22,23*) of ras mutation frequency in lung cancer. K-ras codon 12 may turn out to be only one of several molecular genetic tumor markers that will be useful for lung cancer diagnosis; other potentially useful markers include mutations in other K-ras codons (*22*) and in p53 (*25*). Sensitive assays for detection of molecular tumor markers offer a potentially powerful diagnostic tool for lung cancer.

References

(*1*) Wingo PA, Tong T, Bolden S: Cancer statistics, 1995 [published erratum appears in CA Cancer J Clin 45:127, 1995]. CA Cancer J Clin 45:8-30, 1995

(*2*) Fontana RS: Early diagnosis of lung cancer. Am Rev Respir Dis 116:399-402, 1977

(*3*) Birrer MJ, Brown PH: Application of molecular genetics to the early diagnosis and screening of lung cancer. Cancer Res 52(suppl 9):2658-2664, 1992

(*4*) Szabo E, Birrer MJ, Mulshine JL: Early detection of lung cancer. Semin Oncol 20:374-382, 1993

(*5*) Minna JD, Pass H, Glatstein E, et al: Cancer of the lung. *In* Cancer: Principles and Practice of Oncology (DeVita V Jr, Hellman S, Rosenberg SA, eds), Philadelphia: Lippincott, 1989, pp 591-705

(*6*) Sidransky D, Tokino T, Hamilton SR, et al: Identification of ras oncogene mutations in the stool of patients with curable colorectal tumors. Science 256:102-105, 1992

(*7*) Sidransky D, Von Eschenbach A, Tsai YC, et al: Identification of p53 gene mutations in bladder cancers and urine samples. Science 252:706-709, 1991

(*8*) Bourne HR, Sanders DA, McCormick F: The GTPase superfamily: conserved structure and molecular mechanism. Nature 349:117-127, 1991

(*9*) Rodenhuis S, Slebos RJ: Clinical significance of ras oncogene activation in human lung cancer. Cancer Res (suppl 9)52:2665-2669, 1992

(*10*) Mills NE, Fishman CL, Rom WN, et al: Increased prevalence of K-ras oncogene mutations in lung adenocarcinoma. Cancer Res 55:1444-1447, 1995

(*11*) Merchant RK, Schwartz DA, Helmers RA, et al: Bronchoalveolar lavage cellularity: the distribution in normal volunteers [*see* comment citations in Medline]. Am Rev Respir Dis 146:448-453, 1992

(*12*) Toksoz D, Farr CJ, Marshall CJ: ras gene activation in a minor proportion of the blast population in acute myeloid leukemia. Oncogene 1:409-413, 1987

(*13*) Bashey A, Gill R, Levi S, et al: Mutational activation of the N-ras oncogene assessed in primary clonogenic culture of acute myeloid leukemia (AML): implications for the role of N-ras mutation in AML pathogenesis. Blood 79:981-989, 1992

(*14*) Syvänen AC, Söderlund H, Laaksonen E, et al: N-ras gene mutations in acute myeloid leukemia: accurate detection by solid-phase minisequencing. Int J Cancer 50:713-718, 1992

(*15*) Jacobson DR, Mills NE: A highly sensitive assay for mutant ras genes and its application to the study of presentation and relapse genotypes in acute leukemia. Oncogene 9:553-563, 1994

(*16*) Sugio K, Ishida T, Yokoyama H, et al: ras gene mutations as a prognostic marker in adenocarcinoma of the human lung without lymph node metastasis. Cancer Res 52:2903-2906, 1992

(*17*) Farr CJ, Saiki RK, Erlich HA, et al: Analysis of RAS gene mutations in acute myeloid leukemia by polymerase chain reaction and oligonucleotide probes. Proc Natl Acad Sci U S A 85:1629-1633, 1988

(*18*) Ball ED, Mills LE, Neubauer A, et al: Detection of minimal acute myeloid leukemia cells in bone marrow by probing for mutated ras oncogenes using the polymerase chain reaction. Prog Clin Biol Res 333:499-506, 1990

(*19*) Blin N, Stafford DW: A general method for isolation of high molecular weight DNA from eukaryotes. Nucleic Acids Res 3:2303-2308, 1976

(*20*) Wright DK, Manos MM: Sample preparation from paraffin-embedded tissues. *In* PCR Protocols: a Guide to Methods and Applications (Innis MA, Gelfand DH, Sninsky JJ, et al, eds). San Diego: Academic Press, 1990, pp 153-158

(*21*) Suzuki Y, Orita M, Shiraishi M, et al: Detection of ras gene mutations in human lung cancers by single-strand conformation polymorphism analysis of polymerase chain reaction products. Oncogene 5:1037-1043, 1990

(*22*) Mitsudomi T, Viallet J, Mulshine JL, et al: Mutations of ras genes distinguish a subset of non-small-cell lung cancer cell lines from small-cell lung cancer cell lines. Oncogene 6:1353-1362, 1991

(*23*) Kobayashi T, Tsuda H, Noguchi M, et al: Association of point mutation in c-Ki-ras oncogene in lung adenocarcinoma with particular reference to cytologic subtypes. Cancer 66:289-294, 1990

(*24*) Mao L, Hruban RH, Boyle JO, et al: Detection of oncogene mutations in sputum precedes diagnosis of lung cancer. Cancer Res 54:1634-1637, 1994

(*25*) Greenblatt MS, Bennett WP, Hollstein M, et al: Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. Cancer Res 54:4855-4878, 1994

Notes

Present address: C. L. Fishman, Department of Medicine, Albert Einstein College of Medicine, Bronx, NY.

Supported by Public Health Service grants K12CA0171302 (National Cancer Institute) (N. E. Mills), 5T35DK07421 (National Institute of Diabetes and Digestive and Kidney Diseases) (S. E. Anderson), and T32ES0726702 (National Institute of Environmental Health Sciences) (C. L. Fishman and W. N. Rom), National Institutes of Health, Department of Health and Human Services; and by American Cancer Society institutional grant IRG1435 (D. R. Jacobson).

We thank Dr. Michael Ittman and Dr. Jaishree Jagirdar for assistance in obtaining samples from the pathology archives, Dr. Neil Dubin for assistance with the statistical analysis, Jane Denuto for her efforts in collecting bronchoalveolar lavage fluid samples, Immaculata Kane for technical assistance, and Dr. Joel N. Buxbaum for his support.

Manuscript received November 4, 1994; revised February 27, 1995; accepted May 8, 1995.

PATENT
Docket No.: **19603/481 (CRF D-2472A)**

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s) :	Barany et al.	)	Examiner: B. J. Forman
Serial No. :	09/528,014	)	
Cnfrm. No. :	4478	)	Art Unit: 1634
Filed :	March 17, 2000	)	
For :	COUPLED POLYMERASE CHAIN REACTION-RESTRICTION ENDONUCLEASE DIGESTION-LIGASE DETECTION REACTION PROCESS	)	

NOTIFICATION REGARDING LOSS OF ENTITLEMENT TO SMALL ENTITY STATUS AND PAYMENT OF DEFICIENCY PURSUANT TO 37 CFR §§ 1.27(g)(2) and 1.28(c)

Mail Stop RCE
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

Pursuant to 37 CFR § 1.27(g)(2), applicants hereby notify the U.S. Patent and Trademark Office that the above-identified application is no longer entitled to small entity status.

At the time of filing the above application, small entity status was properly established. Effective as of July 1, 2001, the invention was licensed to a large entity. Therefore, patentees hereby request a change in status from small entity to large entity.

On November 15, 2001, applicants filed an Amendment and a Three-Month Extension of Time and erroneously paid the small entity fee.

On May 19, 2003, applicants filed a Notice of Appeal and a Three-Month Extension of Time and erroneously paid the small entity fee.

Pursuant to 37 CFR § 1.28(c), a check to cover the $1,145.00 deficiency payment is enclosed, for the above-identified application, as shown in the itemization below.

Date	Action – Fees Paid	Amount Paid (Small Entity)	Current Fee (Large Entity)	Amount Owed
11/15/2001	Three-Month Extension of Time	$460.00	$950.00	$490.
05/19/2003	Notice of Appeal	$160.00	$330.00	$170.00
05/19/03	Three-Month Extension of Time	$465.00	$950.00	$485.00
			Current Deficiency Owed	$1,145.00

The Commissioner is authorized to charge any other necessary fees or credit any overpayment to Deposit Account No. 14-1138.

Respectfully submitted,

Date: December 22, 2003

Michael L. Goldman
Registration No. 30,727

NIXON PEABODY LLP
Clinton Square, P.O. Box 31051
Rochester, New York 14603-1051
Telephone: (585) 263-1304
Facsimile: (585) 263-1600

Certificate of Mailing - 37 CFR 1.8(a)

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner for Patents P.O. Box 1450 Alexandria, VA 22313-1450, on the date below.

December 22, 2003 — Jo Ann Whalen

Date — Jo Ann Whalen